Brain Evoked Potential Latencies Optimization for Spatial Auditory Brain-Computer Interface

著者	蔡 振宇
year	2014
その他のタイトル	音の空間情報を利用したブレイン・コンピュータ・インタフェースのための脳波誘発電位の最適判別
学位授与大学	筑波大学 (University of Tsukuba)
学位授与年度	2013
報告番号	12102甲第6860号
URL	http://hdl.handle.net/2241/00123450

Brain Evoked Potential Latencies Optimization for Spatial Auditory Brain-Computer Interface

Graduate School of Systems and Information Engineering

University of Tsukuba

March 2014

ZHENYU CAI

Declaration of Authorship

I, Zhenyu Cai, declare that this thesis titled, 'Brain Evoked Potential Latencies Optimization for Spatial Auditory Brain-Computer Interface' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

UNIVERSITY OF TSUKUBA

Abstract

Graduate School of Systems and Information Engineering

Doctor of Computer Science

Brain Evoked Potential Latencies Optimization for Spatial Auditory Brain-Computer Interface

by Zhenyu CAI

The thesis presents a report on the auditory BCI. First, I utilized a novel auditory BCI paradigm based on combined sound timbre and horizontal plane spatial locations as informative cues. The presented concept is based on responses to eight–directional audio stimuli with various tonal and environmental sound stimuli. The approach is based on a monitoring of brain electrical activity by means of the electroencephalogram (a short name called "EEG"). The first achievement discussed in this thesis is a BCI analysis based on optimization of electrode locations on the scalp for further classification accuracies improvement. Based on this analysis results, it allows us decrease the electrodes number from 64 to 10. Next, I propose a methodology for finding and optimizing brain evoked response latencies in the $P300$ range in order later to classify them correctly and to elucidate the subject's chosen targets or ignored non-targets. To accomplish the above, I propose an approach based on an analysis of variance for feature selection. Such results also unsatisfactory as regards as a successful BCI system application. Third, I design a new auditory spatial auditory BCI paradigm in which the ERP shape differences at early latencies are employed to enhance the $P300$ responses in an oddball experimental setting. The concept relies on the recent results in auditory neuroscience showing a possibility to differentiate early anterior contralateral responses to attended spatial sources. Contemporary stimuli–driven BCI paradigms benefit mostly from the $P300$ ERP latencies in so called "aha-response" settings. I show the further enhancement of the classification results in spatial auditory paradigms by incorporating the $N200$ latencies, which differentiate the brain responses to lateral, in relation to the subject head, sound locations in the auditory space. I also found that the early brain responses elucidate which direction, front or rear loudspeaker source, subject attended. Such results reveal that those spatial auditory ERPs boost classification results of the BCI application. The BCI experiments with the multi-command BCI prototype support our research hypothesis with the higher classification results and the improved information-transfer-rates.

Acknowledgements

I would like to thank Prof. Shoji Makino, Prof. Tomasz M. Rutkowski and Prof. Takeshi Yamada for their many valuable suggestions and constant support during this research. I also would like to thank Prof. Hiroyuki Kudo, Prof. Ko Sakai and Prof. Hotaka Takizawa for their many valuable suggestions and comments. Prof. Shoji Makino, my advisor, for supporting me to various seminars and conferences, and for many valuable suggestions and comments in this research. My co-advisor, Prof. Tomasz M. Rutkowski, who contributed a lot to this research, and who helped me so many times with my experiments.

This research was supported in part by the Strategic Information and Communications R D Promotion Program no. 121803027 of The Ministry of Internal Affairs and Communication in Japan, and by KAKENHI, the Japan Society for the Promotion of Science, grant no. 12010738 and 24700154. I also acknowledge the technical support of YAMAHA Sound and IT Development Division in Hamamatsu, Japan.

Contents

Declaration of Authorship i

Abstract ii

Acknowledgements iv

List of Figures viii

List of Tables xii

1 Introduction 1
- 1.1 Objectives and Scope . . . 3
  - 1.1.1 Aims of the Thesis . . . 4
  - 1.1.2 Organization of the Thesis . . . 4
    - 1.1.2.1 Chapter 2: Reviews Existing Technology . . . 5
    - 1.1.2.2 Chapter 3: Electrodes Position Optimization . . . 5
    - 1.1.2.3 Chapter 4: $P300$ Optimization for Spatial auditory BCI . . . 5
    - 1.1.2.4 Chapter 5: Utilization of $N200$ and $P300$ for Spatial Auditory BCI Enhancement . . . 6
    - 1.1.2.5 Chapter 6: ERP Responses to Front–Back to the Head Stimuli Distinction Support . . . 6

2 Reviews of Existing Technology 8
- 2.1 Recording brain activity . . . 8
  - 2.1.1 Invasive brain Recording . . . 8
  - 2.1.2 Recording based on Hemodynamic Response . . . 9
  - 2.1.3 Recording based on Electromagnetic waves . . . 12
  - 2.1.4 Summary of Brain Signal Monitoring . . . 12
  - 2.1.5 Imagery BCI based on EEG . . . 13
    - 2.1.5.1 Imagery BCI Controlled Wheelchair . . . 14
  - 2.1.6 Stimuli-driven BCI based on EEG . . . 16
    - 2.1.6.1 Event Related Potential . . . 16
    - 2.1.6.2 P300 Response . . . 17

2.1.7 Review of stimuli-driven BCI 17
2.1.7.1 Visual Keyboard Speller BCI 18

3 Electrodes Position Optimization 20
3.1 Introduction 20
3.1.1 Methods 21
3.1.1.1 EEG Recording System 22
3.1.1.2 Stimulus 23
3.1.1.3 Spatial Hearing 24
3.1.1.4 EEG Experiment Description 24
3.1.2 Analysis 26
3.1.2.1 EEG Preprocessing 26
3.1.2.2 Electrodes and ERP Features Selection for Classification 27
3.1.3 Results 30
3.1.4 Conclusions 33

4 *P*300 Optimization for Spatial aBCI 36
4.0.5 Introduction 36
4.0.6 Method 37
4.1 Psychophysical Experiment 38
4.1.1 The Offline aBCI Experiment Protocol 39
4.1.2 EEG Acquisition 40
4.1.3 EEG Response Analysis 40
4.1.3.1 EEG Preprocessing 42
4.1.3.2 ERP Feature Extraction Using ANOVA of the ERP Latencies 42
4.1.3.3 The Offline ERP Classification in the aBCI Paradigm 44
4.2 Results 47
4.2.1 Analysis of aBCI Results with ITR and Classification Accuracies. 47
4.2.2 The ITR and Classification Accuracy Results from the *P*300 ERP Range Latencies in the Single Channel Setting of *Target* versus *Non-target* 47
4.2.3 The ITR Results from the *P*300 ERP Range Latencies from the Averaged Eight Trials in the Setting of *Target* versus *Non–target* 48
4.3 Discussion and Conclusions 48

5 Utilization of *N*200 and *P*300 for Spatial aBCI Enhancement 55
5.1 Introduction 55
5.2 Method 57
5.2.1 The Offine saBCI Experimental Protocol 58
5.2.1.1 The Analysis of ERP Responses in Offine BCI Paradigm 59

5.2.2 EEG Preprocessing . . . 59
5.2.3 The Optimization of the EEG Electrode Locations and ERP Features Extraction . . . 60
5.2.4 The Offline saBCI Classification . . . 64
5.3 Results . . . 68
5.3.1 The Classification Results from the Combined $N200$ and $P300$ ERP Latencies in the Classical *target* vs. *non–target* Setting . . . 68
5.3.2 The Classification Results from the new $N2apc$ ERP Feature in the Ipsilateral vs. Contralateral Settings . . . 68
5.3.3 Analysis of Information Transfer Rate Improvement Results 69
5.4 Conclusions . . . 69

6 ERP Responses to Front–Back to the Head Stimuli Distinction Support 73
6.1 Introduction . . . 73
6.2 Methods . . . 74
6.2.1 The Offline saBCI Experimental Protocol . . . 76
6.3 The Analysis of ERP Responses in Offline BCI Paradigm . . . 77
6.3.1 EEG Preprocessing . . . 78
6.3.2 The Optimization of the EEG Electrode Locations and ERP Feature Extraction . . . 78
6.3.3 The Offline saBCI Classification . . . 81
6.4 Results . . . 85
6.4.1 The Classification Results from the $P300$ ERP Latencies in the Classical Oddball Paradigm Setting . . . 85
6.4.2 The Classification Results from the $N2ac$ ERP Feature in the Ipsilateral vs. Contralateral Settings . . . 86
6.4.3 The classification Results of the $N2fr$ ERP Feature Extraction Method . . . 86
6.4.4 Analysis of Information Transfer Rate Improvement Results 86
6.5 Conclusions . . . 86

7 Conclusion of thesis 91
7.1 Summary of thesis . . . 91
7.2 Discussion . . . 93

Bibliography 94

Author's Paper List 100

List of Figures

2.1 fMRI recording equipment (upper figure) and fMIR signals (lower figure) The figures source: Bergen fMRI Group, Department of Biological and Medical Psychology, University of Bergen http://fmri.uib.no/ 10

2.2 fNIR recording equipment (The figure source: SING YIP TECHNOLOGY (HK) CO., LIMITED) http://www.biopac.com/fnir-optical-brain-imaging-hemodynamic-response 11

2.3 MEG recording equipment Figure source: The National Institute of Mental Health (NIMH) http://www.nimh.nih.gov/index.shtml . 13

2.4 64 channels EEG data were recorded by Biosemi 14

2.5 With the RIKEN's brain controlled wheelchair, the user continuously controls the velocity of the wheelchair by imagining left hand, right hand or both feet movements (Figure obtained permission) . . 15

2.6 In the $P300$ speller by Farwell and Donchin, items are presented on a 6 by 6 matrix. Rows and columns are flashed in a random sequence, eliciting a $P300$ signal 300 ms after the key the user wants to select has been flashed. The figure source: http://www.bbci.de/competition/ii/ 19

3.1 Spatial auditory BCI paradigm concept with octagonal loudspeaker arrangement. 22

3.2 Spatial auditory BCI open acoustic filed. 23

3.3 User training interface for spatial auditory BCI experiment. 25

3.4 MAX/MSP path managing the spatial BCI Experiment. 25

3.5 BIOSEMI system with 64 electrodes. 27

3.6 Results of ERP $P300$ response for frontal (upper panel) and rear (lower panel) speakers confirming the feasibility of the proposed approach. The zero stands for stimuli onsets. The blue/dashed lines depict *non-target* (no $P300$ response after 300 ms) responses and red/solid traces visualize attended spatial *targets* (obvious positive EEG response deflections after in $300-500$ ms range). 28

3.7 ROC analysis results of a good channel candidate (col1: black/upper trace) discriminating and a "chance level" one for not channel selection (col2: red/lower trace). 30

3.8 Results of classical LDA application to binary classification of a spatial tonal 440 Hz stimuli applied to all electrodes for a single subject in six crossvalidation trials are visualized in the left column while the best resulting electrodes are presented in the right column for each subject. All graphs have the same scaling.. 31

3.9 Results of the proposed ROC analysis based channel selection for all six subjects and two stimuli cases revealing the temporal and parietal scalp regions as best candidates for spatial stimuli $P300$ responses identification. 32
3.10 Percentage of correctly classified spatial audio BCI $P300$ responses with LDA classifiers derived from three experimental sessions for each subject (blue/solid lines) and based on ten best electrode results (red/dotted lines). Chance level is 50%. 34
3.11 Comparison of classification responses to frontal (top panel) and rear/back (bottom panel) loudspeakers stimuli directions for six subjects and two stimuli conditions. The results confirm only slight subjects' preferences to frontal stimuli directions except of subject #6 who had better results for rear sound directions.. 35

4.1 Spatial auditory BCI paradigm concept with eight loudspeakers in the upper part of the figure. The lower graph visualizes the stimulus presentation concept in the time domain. Each stimulus is presented for 30 ms with 170-ms silent breaks, so the ISI is set to 200 ms. . . 39
4.2 Eight electrodes on the scalp location (see the blue shadow) 41
4.3 Spatial auditory BCI paradigm concept with eight loudspeakers in the upper part of the figure. The lower graph visualizes the stimulus presentation concept in the time domain. Each stimulus is presented for 30 ms with 170-ms silent breaks, so the ISI is set to 200 ms. . . 43
4.4 Grand mean average auditory evoked responses to spatial white noise stimuli of the ten subjects from the eight electrodes plotted separately in each row of the panels. The top panel shows the grand mean averaged response to the *targets*. The middle panel presents the grand mean averaged responses to *non-targets*. The bottom panel depicts the p values from the ANOVA for the eight electrodes separately 45
4.5 Grand mean average auditory evoked responses to spatial pink noise stimuli of the ten subjects from the eight electrodes plotted separately in each row of the panels. The top panel shows the grand mean averaged response to the *targets*. The middle panel presents the grand mean averaged responses to *non-targets*. The bottom panel depicts the p values from the ANOVA for the eight electrodes separately 46

5.1 The novel N2apc paradigm based on spatial sound stimuli. 56
5.2 Six electrodes for testing N2ac on the scalp location (see blue shadow) 61
5.3 The grand mean averaged ERP responses of the seven subjects. The solid lines depict *targets* and the dashed ones *non-targets*. The red color indicates ipsilateral and blue one the contralateral responses. The differences between *targets* and *non-targets* are obvious after 300 ms (the so called "aha" or $P300$ response), while the lateral directions can be identified in $N200$ latency area. 62

5.4 The grand mean averaged ERP for the all seven subjects and all electrodes calculated together, while plotted separately for *target* (solid red line) and non–target (dashed blue line) responses. The significant differences between the both responses can be found, as visualized by the color bar with p–values of t–test results (statistical significance for $p < 0.05$) in the bottom part in the above panel, can be found around 200 ms ($N200$ response latency) and after 300 ms ($P300$ response latency). 63
5.5 ERP to pink noise stimuli grand mean averages for all subjects and the six electrodes plotted separately in each panel. The solid red lines represent the ipsilateral to *target* responses and the dashed blue lines to the contralateral ones. The color bars at the bottom of each panel show the *t-test* resulting $p-$values. 65
5.6 ERP to white noise stimuli grand mean averages for all subjects and the six electrodes plotted separately in each panel. The solid red lines represent the ipsilateral to *target* responses and the dashed blue lines to the contralateral ones. The color bars at the bottom of each panel show the *t-test* resulting $p-$ values. 66

6.1 The novel front and rear to the subject head's auditory sources localization paradigm based on spatial sound stimuli is depicted in the top panel. The bottom panel presents our stimulus presentation concept illustrated in the time domain. Each stimulus has been presented in our experiments for 30 ms with 270 ms silent breaks with the respective inter–stimulus–interval (ISI) of 300 ms. 75
6.2 The grand mean averaged ERP responses of the seven subjects. The upper panel present the grand mean averaged ERP responses to pink–noise. The lower panel depicts respective results obtained with white–noise stimulus. The red lines represent *targets* and the blue ones non-targets. All results are presented together with standard error bars. The differences between *targets* and non-targets are obvious in the range of 300 $\sim$ 600 ms (the so called "aha"– or $P300$ response). 79
6.3 Grand mean averaged ERP responses of the all seven subjects and the six electrodes are plotted separately in the two top panels. The first top panel shows *target* and a second from the top the *non–target* averaged responses, respectively. The significant differences between the both responses can be found, as visualized by color coding of the p-values obtained from *ANOVA–test* (statistical significance at $p < 0.05$) in the third from the top panel. The bottom panels presents *Kolmogorov–Smirnov test*. The EEG electrodes $F5, F6, C3, C4, P5$ and $P6$ were used in the experiments. 82

6.4 Grand mean average ERP responses of all seven subjects and the six electrodes plotted separately in each panel in *contralateral* (top panel) vs. *ipsilateral* (second from the top panel) stimulus direction presentation settings. The significant differences between the both responses can be found, as visualized by the color with p-values of *ANOVA–test* and *KS–test* results (statistical significance for $p < 0.05$) in the third panel and fourth panel, respectively. EEG electrodes $F5, F6, C3, C4, P5$ and $P6$ were used in the experiment. 83

6.5 Grand mean averaged ERP responses of all subjects and the six electrodes plotted separately in each panel for *the frontal–* (top panel) and *the rear–loudspeaker* (second from the top panel) ERP responses. The significant differences between the both responses can be found, as visualized by the color with p-values of *ANOVA–test* and *KS–test* results (statistical significance at $p < 0.05$) in the third and fourth from the top panelz respectively. EEG electrodes $F5, F6, C3, C4, P5$ and $P6$ were used in the experiment. 84

List of Tables

4.1 The spatial sound psychophysical experiment results. The response time delays and instructed directions accuracies are presented in the form of mean values with standard deviations (STD) respectively. 50

4.2 The confusion matrix results from the psychophysical experiment averaged for all the subjects for pink– and white–noise stimulus respectively. 50

4.3 The offline aBCI interfacing results based on features drawn from non–averaged trials in the form of ITR scores obtained as in equations (4.1) and (4.2). I compare the traditional *all ERP* and the proposed "hand–picked" only latencies. 51

4.4 The classification results for ERP latencies in $P300$ responses for *target* vs. *non–target* paradigm. The classification results of two feature sets (*allERP* responses and $P300$ responses) are compared. The classification improvement comparing the conventional all ERP latency with the proposed $P300$ response) is summarized in the right column. 52

4.5 The offline aBCI interfacing results based on features drawn from the averaged eight trials in the form of ITR scores obtained as in equations (4.1) and (4.2). I compare the traditional *whole–ERP* and the proposed "hand–picked" only latencies. 53

4.6 The classification results for ERP latencies in $P300$ responses for the mean of 8 *targets* vs. average of 8 *non–targets* paradigm. The classification results for two feature sets (*allERP* responses and $P300$ responses) are compared. The classification improvement comparing the conventional all ERP latency with the proposed $P300$ response) is summarized in the right column. 54

5.1 The classification results for ERP latencies in $N200$ and $P300$ responses for *target* vs. *non–target* paradigm. The three feature sets ($N200$, $P300$ and $N200/P300$ latencies combined) classification results are compared. The classification improvement comparing the classical $P300$ latency only with the proposed combination of $N200/P300$) is summarized in the right column. 71

5.2 The classification results for the proposed method using $N2apc$ response to support the saBCI compared with the conventional method. 71

5.3 The ITR for the three ERP interval related classification approaches using $N200$ or $P300$ only, and the combined $N200/P300$ together. . 72

5.4 The ITR for the proposed method using $N2apc$ response to support the saBCI classification rates. 72

6.1 The classification results for ERP latencies in $P300$ responses for *target* vs. *non–target* paradigm. The three sets (whole ERP, $P300$ latencies optimized by ANOVA mehtod and KS method) classification results are compared. 87
6.2 The classification results for ERP latencies in $N200$ responses for *ipsilateral* vs. *contralateral* paradigm. The three feature sets (whole ERP (0 ms – 700 ms), $N200$ latencies optimized by ANOVA method and KS method) classification results are compared. . . . 88
6.3 The classification results for ERP latencies in $N200$ responses for *targets from frontal loudspeaker* vs. *targets from rear loudspeaker* paradigm. The three feature sets (whole ERP (0 ms – 700 ms), $N200$ latencies optimized by ANOVA method and KS method) classification results are compared. 89
6.4 The ITR of *target* vs. *non-target*, for the three ERP sets related classification approaches using $wholeERP$, $P300$ extracted with ANOVA and KS. 89
6.5 The ITR of N2ac, for the three ERP sets related classification approaches using $wholeERP$, $N200$ extracted with ANOVA and KS. 90
6.6 The ITR of *targets* from front loudspeaker and *targets* from rear loudspeaker, for the three ERP sets related classification approaches using $wholeERP$, $N200$ extracted with ANOVA and KS. 90

Chapter 1

Introduction

Human need to communicate and interact with the environment, this is a human basic life needs. Severe motor disabilities can limit one's ability to communicate, especially for patients suffering from amyotrophic lateral sclerosis (ALS), severe cerebral palsy, head trauma, multiple sclerosis, and muscular dystrophies who are incapable of conveying their intentions (locked-in syndrome) to the external environment. Considering that the ALS is the second major neurodegenerative disease with an incidence of $5/100,000$ per year (twice more important than Parkinson disease incidence) and that this disease occurs in adulthood. Such patients can not meet the basic needs. In the past several decades, A novel communication techniques which called BCI (BCI: brain computer interface) is developed. Such this novel techniques is designed to establish a communication link between the human brain and a computer [1]. It is use electric, magnetic or hemodynamic brain activity to control external devices such as robot, computers, wheelchairs, etc. A BCI does not depend on muscle or peripheral nervous system activity. In particular, a BCI could help patients suffering from amyotrophic lateral sclerosis (ALS) to communicate or to complete various daily tasks, including controlling a computer or typing messages on a virtual keyboard. BCI researchers attempt to develop new communication pathways for ALS paints using their brain signals, also focus on developing potential applications in rehabilitation and multimedia communication. Most of current BCIs have three major functions [2] :

- measure signals from the human brain;
- methods and algorithms for recognize brain states/ intentions from these signals;

- methodology and algorithms for mapping human brain activity to intended behavior or action.

Several different technologies available to record the brain activity (see chapter 2). Such as electroencephalography (EEG), magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI), and functional Near-Infrared Imaging (fNIR). However, MEG and fMRI are expensive and gigantic, and fMRI show longtime constants in that they do not measure neural activity directly which relying instead on the hemodynamic coupling between neural activity and regional changes in blood flow. Only fNIR imaging and EEG can be used for a BCI. And yet, NIR is a relatively new method, it is not as popular as EEG in BCI studies. As this result, the majority of BCIs utilize EEG signals [1], [2].

Most of BCI studies are based on visual evoked potential (VEP) (described in chapter 2). VEP use in BCI was introduced by [3] and this study also became a major focus in BCI research. They describes a BCI that is depended on auditory evoked potential(AEP) (see chapter 2), which may extend the study on VEP in several important aspects. As I mentioned before, auditory BCI (aBCI) has become a hot topic of great interest in computational auditory neuroscience. The aBCI utilize human auditory pathway responses and allows users to operate external devices more quickly and simply, based on auditory evoked responses to sound stimuli. In an former research [4], it is describe that 61 people suffering from mid-stage ALS were included with ALSFRS-R ratings ranging from 18 to 33. The author found that a larger portion of their population suffered from auditory (42%) than visual (24%) deficits. Recently, the other study reported that a patient with ALS through the late stages of the disease, BCI may only be possible with the auditory pathways [5]. Indeed, Two researches show that the visual BCIs may not suitable when eye-gaze control is limited [6], [7]. For this reasons, the novel and interest in auditory BCI has grown in recent years.

Auditory stimulation involves an oddball paradigm where the subject is attend to stimuli that differ from each other on some property. The subject is required to focus on one of them. A binary BCI generally using such paradigms which utilize tones with various of pitch [8] , [9]. In paper [8], authors presented two sequences of *target* and *non-target* tones to the subject. The subject's ears listened simultaneously a sequence with a different inter stimulus interval. The subjects required to pay attention on either one of stimulus with counting the number of them. The time samples for left and right *non-target* tones were taken from the same four seconds of EEG signal, averaged and subsequently concatenated. For the different inter stimulus interval, Event-related potential (ERP) (see chapter 2) in response to the left channel would average out on right channel. Such feature was used for later binary classification (*target* vs. *non-target*). Even through the classification rate have a big different between of them, the results showed that it is possible for BCI using auditory ERP as a feature.

Several researches show that auditory ERP based on spatial sound stimuli can improve BCI performance [10], [11]. In offline BCI studies, additional spatial information to an auditory stimulus was helpful to the recognition, reaction times and also including classification accuracy [12]. Such paradigms could even be used to extend visual BCIs to improve performance during the late stages ALS [13].

1.1 Objectives and Scope

A common feature between all of BCIs is that, since the recorded brain signals are very nosiy and a low classification accuracy. In order to solve the low classification accuracy problem and make it is possible to be used in the online BCI applications. This is the challenge I address in this thesis.

Application of evoke-BCIs are reported in the literate, mostly of them are visual BCIs. However, a auditory BCI have three merits comparing with visual BCIs is:

- a less mentally involving sense;
- to operate exteral devices more quickly and simply;
- it is could be used by the late stages ALS patients;

In the thesis, I propose to utilize a concept of a spatial auditory stimuli. This is a very interesting concept since, in contrary to the contemporary visual modality BCI, it utilizes audition. During the BCI experiments, the subject intentionally directs his attention to different locations in surround sound environment. A sound direction which the subject attends evokes $N200$, $P300$ response which is next used to generate a BCI command. In experiments described in the thesis I utilize octagonal loudspeakers setup. The setup and obtained results are a step forward in comparison to contemporary reports since I am able to classify also the $N200$, $P300$ responses from rear-side-loudspeakers, which was not reported till now.

1.1.1 Aims of the Thesis

As mentioned above, to be successful. The novel auditory BCI has to achieved the following purpose: fully octagonal surround sound, high classification accuracy.

- *fully octagonal surround sound*: The spatial aBCI concept is founded on a basic feature of the human auditory pathway, which is very sensitive to the location of changing spatial auditory sources. The auditory pathway also has a very good temporal resolution, which is an additional feature we would like to utilize in the spatial aBCI design. This will make it possible to reduce inter-stimuli intervals (ISI) of the presented sounds in comparison with vision-based applications. Contemporary applications have thus far failed to use rear-to-the-head loudspeakers, as postulated as an optimal setting yet still not fully realized in Schreuder et al [12].
- *Hight classification accuracy*: we discuss a novel auditory BCI paradigm with the support of the ERP component ($P300$, $N200$), evoked by the expected/instructed *targets* [14]. Our hypothesis is that a significant ERP response will be found when subjects attend to the *target* direction and ignore the *non-targets*.

A BCI is based on a neuroscience research with support of signal processing and machine learning algorithms. Our goal in this work is conducting research on novel algorithms and a design of new BCI experimental paradigms. This requires the novel algorithms able to improve the classification accuracy, and using spatial sound stimuli with fully octagonal surround sound could make a multi-command auditory BCI.

1.1.2 Organization of the Thesis

This thesis is composed of six chapters. A description of each chapter, excluding the present chapter 1, is given as follows:

1.1.2.1 Chapter 2: Reviews Existing Technology

In this chapter, several ways of brain recording and BCI system are described. This chapter provides the reader with a theoretical background and an extensive BCI technology, It helps to understand the four studies described in Chapter 3, Chapter 4, Chapter 5 and Chapter 6.

1.1.2.2 Chapter 3: Electrodes Position Optimization

Our proposed method 1

In this chapter, it introduces a novel auditory BCI paradigm based on combined sound timbre and horizontal plane spatial locations as informative cues. The presented concept is based on responses to eight–directional audio stimuli with various tonal and environmental sound stimuli. The approach is based on a monitoring of brain electrical activity by means of the electroencephalogram (EEG). The previously developed by the authors spatial auditory stimulus is extended to varying in timbre sound stimuli which feature helps the subjects to attend to the *targets*. The main achievement discussed in chapter 3 is an offline BCI analysis based on an optimization of electrode locations on the scalp and evoked response latency for further classification results improvement. The so developed new BCI paradigm is more user-friendly and it leads to better results comparing to previously utilized simple tonal or steady-state stimuli.

1.1.2.3 Chapter 4: $P300$ Optimization for Spatial auditory BCI

Our proposed mehtod 2

In this chapter, we propose a novel method for the extraction of discriminative features in electroencephalography (EEG) evoked potential latency. Based on our offline results, we present evidence indicating that a full surround (eight–directions) sound auditory BCI paradigm has potential for an online application. The auditory spatial BCI concept is based on chapter 3, which employs a loudspeaker array in an octagonal horizontal plane. The stimuli presented to the subjects vary in frequency and timbre. To capture brain responses, we utilize an eight-channel EEG system. I propose a methodology for finding and optimizing evoked response latencies in the $P300$ range in order later to classify them correctly and to elucidate the subject's chosen *targets* or ignored *non-targets*. To accomplish the above, we propose an approach based on an analysis of variance for feature selection. Finally, we identify the subjects intended commands with a naive Bayesian classifier for sorting the final responses. The results obtained with ten subjects in offline B-CI experiments support our research hypothesis by providing higher classification results and an improved information transfer rate compared with state-of-the-art solutions.

1.1.2.4 Chapter 5: Utilization of $N200$ and $P300$ for Spatial Auditory BCI Enhancement

Our proposed mehtod 3

Chapter 5 presents our results obtained with a new auditory spatial localization based BCI paradigm in which the ERP shape differences at early latencies are employed to enhance the $P300$ responses in an oddball experimental setting. The concept relies on the recent results in auditory neuroscience showing a possibility to differentiate early anterior contralateral responses to attended spatial sources. Chapter 4 BCI paradigms benefit mostly from the $P300$ ERP latencies. I show the further enhancement of the classification results in spatial auditory paradigms by incorporating the $N200$ latencies, which differentiate the brain responses to lateral, in relation to the subject head, sound locations in the auditory space. The results reveal that those early spatial auditory ERPs boost online classification results of the BCI application. The online BCI experiments with the multi-command BCI prototype support our research hypothesis could improve classification accuracies and information-transfer-rates.

1.1.2.5 Chapter 6: ERP Responses to Front–Back to the Head Stimuli Distinction Support

Our proposed method 4

Chapter 4 shows a possibility to differentiate early anterior contralateral responses to attended spatial sources. I also found that the early brain responses elucidate which direction, front or rear loudspeaker source, subject attended. I show the further enhancement of the classification results in spatial auditory paradigm, in which I incorporate the $N200$ latencies. The results reveal that those early s-patial auditory ERPs boost offline classification results of the BCI application. The offline BCI experiments with the multi-command BCI prototype support our research hypothesis with the higher classification results and the improved information–transfer–rates.

Chapter 2

Reviews of Existing Technology

In this chapter I introduce what are the different technologies available to record (Section 2.1) the brain activity. Then, in Section 2.2 I review two Representative stimuli-driven-BCI applications.

2.1 Recording brain activity

The first step toward a BCI is recording the activity of the human brain. This can be done invasively by surgically implanting electrodes in the brain, or non-invasively. In this section I will review different brain activity recording technologies which I mentioned in Chapter 1.

2.1.1 Invasive brain Recording

Biologists can measure the potential at different parts of a single neuron in a culture. Recording neuron activity in a living brain is possible using surgically implanted micro-electrodes arrays, although it is no longer a single neuron recording but the activity of groups of neurons. Monkeys with brain implants have been reported to brain-control the displacement of a cursor on a screen or to control the motion of a robotic arm. Surgical implantation of electrodes is still consider too dangerous to be carried on humans brain [15]. However, this resolution is sufficient to show physiological processes at the cellular level. Clinical approaches sometimes allow invasive recordings to be taken from the human brain, mainly in patients with epilepsy or with movement disorders, and such recordings can sample neural activity at spatial scales ranging from single cells to distributed cell assemblies. Two typical representative of invasive brain recording are single unit/neuron recordings and electrocorticography.

- *single unit/neuron recordings*: The use of an electrode to record the electrophysiological activity is(action potentials) from a single neuron. An electrode introduced into the brain of a living animal detect electrical activity that is generated by the neurons adjacent to the electrode tip. If the electrode is a micro–electrode, with a tip size of 3 to 10 micrometers, the electrode often isolate the activity of a single neuron.
- *ElectroCorticoGraphy (ECoG)*: ECoG is an invasive procedure Because a craniotomy (a surgical incision into the skull) is required to implant the electrode grid. The electrodes are placed directly on the exposed surface of the brain to record electrical activity from the cerebral cortex.

2.1.2 Recording based on Hemodynamic Response

The typical hemodynamic response method include Functional Magnetic Resonance Imaging (fMRI) and Near-Infrared Imaging.

Functional Magnetic Resonance Imaging (fMRI):

- fMRI is a relatively recent imaging technique that aims to determine the neuro–biological correlate of behavior by identifying the brain regions that become "active" during the performance of specific tasks in vivo [16]. This recording method is a functional neuro–imaging procedure using magnetic resonance imaging technology that measures brain activity by detecting associated changes in blood flow. The technique is based upon the different magnetic susceptibilities of the iron in oxygenated and deoxygenated hemoglobin (see Figure 2.1). Oxygenated blood is diamagnetic and possesses a small magnetic susceptibility, while deoxygenation of hemoglobin produces deoxyhemoglobin, which is a significantly more paramagnetic species of iron. Blood Oxygenation Level Dependent (BOLD) measurements measure local variation in the relaxation time caused by variations in the local concentration of deoxygenated blood. It has become the diagnostic method of choice for investigating how a normal, diseased or injured brain are working. The spatial resolution can be sub-millimeter with temporal resolutions on the order of seconds. The ability to measure solitary neural events is not yet possible but improvements in sensitivity have been made steadily over the past decades.

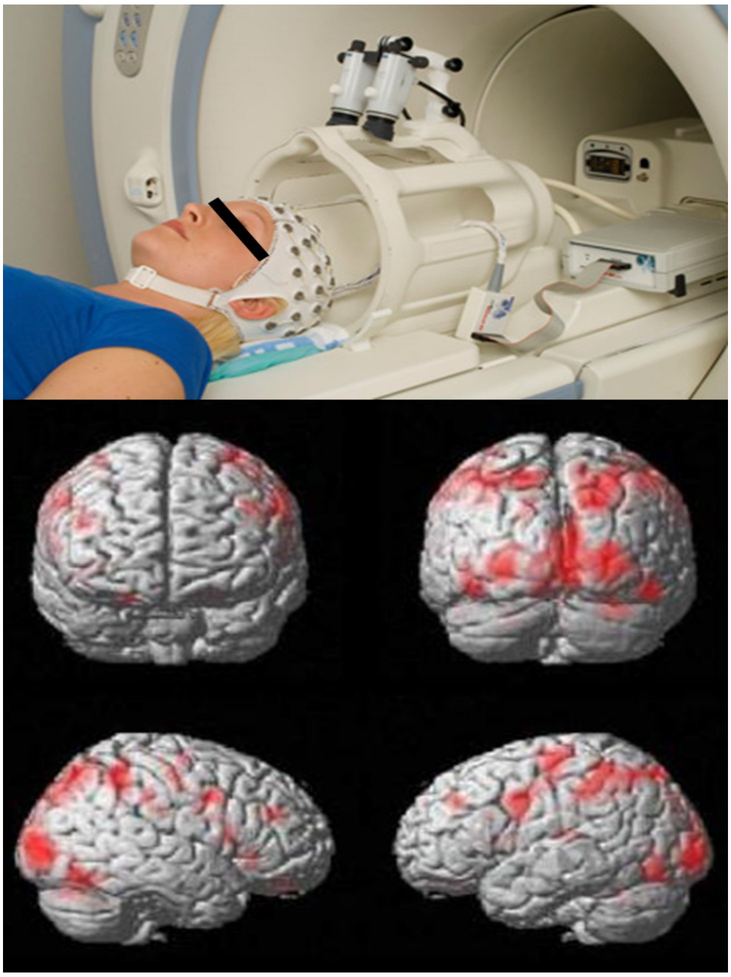

FIGURE 2.1: fMRI recording equipment (upper figure) and fMIR signals (lower figure) The figures source: Bergen fMRI Group, Department of Biological and Medical Psychology, University of Bergen http://fmri.uib.no/

Functional Near-Infrared Imaging (fNIR):

- fNIR is a relatively novel technology based upon the notion that the optical properties of tissue (including absorption and scattering) change when the tissue is active (see figure 2.2). Two types of signals can be recorded: fast scattering signals, presumably due to neuronal activity and slow absorption signals, related to changes in the concentration of oxy- and deoxy-hemoglobin [17]. However, fNIR lacks the spatial resolution of fMRI and cannot accurately measure deep brain activity. The fast fNIR signal is measured as an event-related optical signal (EROS). The spatial localization of fast and slow fNIR measurements both correspond to the BOLD fMRI signal. The latency in the slow (hemodynamic) signal roughly corresponds to that for the BOLD fMRI response. The major limitation of optical methods (both fast and slow signals) is their penetration (max: approximately 3 cm from head to surface), which makes it impossible to measure brain structures such as the hippocampus or the thalamus, especially if they are surrounded by light-reflecting white matter. However, the vast majority of the cortical surface is accessible to the measurements. The technology is relatively simple and portable, and may serve a sort of portable, very rough equivalent of fMRI, which may supplement or substitute for some EEG measures.

Figure 2.2: fNIR recording equipment (The figure source: SING YIP TECHNOLOGY (HK) CO., LIMITED) http://www.biopac.com/fnir-optical-brain-imaging-hemodynamic-response

2.1.3 Recording based on Electromagnetic waves

The currents generated by an individual neuron are too tiny to be recorded non-invasively, however excitatory neurons in the cortex all have their axon parallel one to another and grouped in redundant populations called macro-columns which act as macroscopic sources of electromagnetic waves that can be recorded non-invasively.

Magnetoencephalography (MEG)

- MEG is an imaging technique used to map brain activity by recording magnetic fields produced by electrical currents occurring naturally in the brain, using very sensitive magneto meters [18], [19]. Because of the low strength of these signals and the high level of interference in the atmosphere, MEG has traditionally been performed inside rooms designed to shield against all electrical signals and magnetic field fluctuations (see Figure 2.3).

Electroencephalography (EEG)

- EEG is the recording of electrical activity along the scalp produced by the firing of neurons within the human brain [14], [20]. The recording is obtained by placing electrodes on the scalp with a conductive gel or dry electrodes without gel. The number of electrodes depends on the application, from a few to 256, and they can be mounted on a cap for convenience of use. The electric signal recorded is of the order of few microvolt, hence must be amplified and filtered before acquisition by a computer (see Figure 2.4). The electronic hardware used to amplify, filter and digitize the EEG signal is of the size and weight of a book; it is easily transportable and relatively affordable. Spatial resolution is on the order of centimeters while the time of response to a stimulus is on the order of 100 milliseconds.

2.1.4 Summary of Brain Signal Monitoring

The six methods presented above. Only fNIRs imaging and EEG can be used for a BCI: MEG and fMRI equipments are too expensive and cumbersome, and invasive methods are not safe enough yet. However, as fNIRs is a relatively new method, it is not as popular as EEG in BCI studies.

FIGURE 2.3: MEG recording equipment Figure source: The National Institute of Mental Health (NIMH) http://www.nimh.nih.gov/index.shtml

2.1.5 Imagery BCI based on EEG

Imagery BCI is an independent mode paradigm in which only subject brain intentional imagery patterns are classified to generate interfacing commands (imagination/planning of a right or left hand movement, etc.). An electric wheelchair driven through a brain-computer interface (BCI), developed by RIKEN Brain-TOYOTA Collaboration Center in Japan. The system is described in the next section.

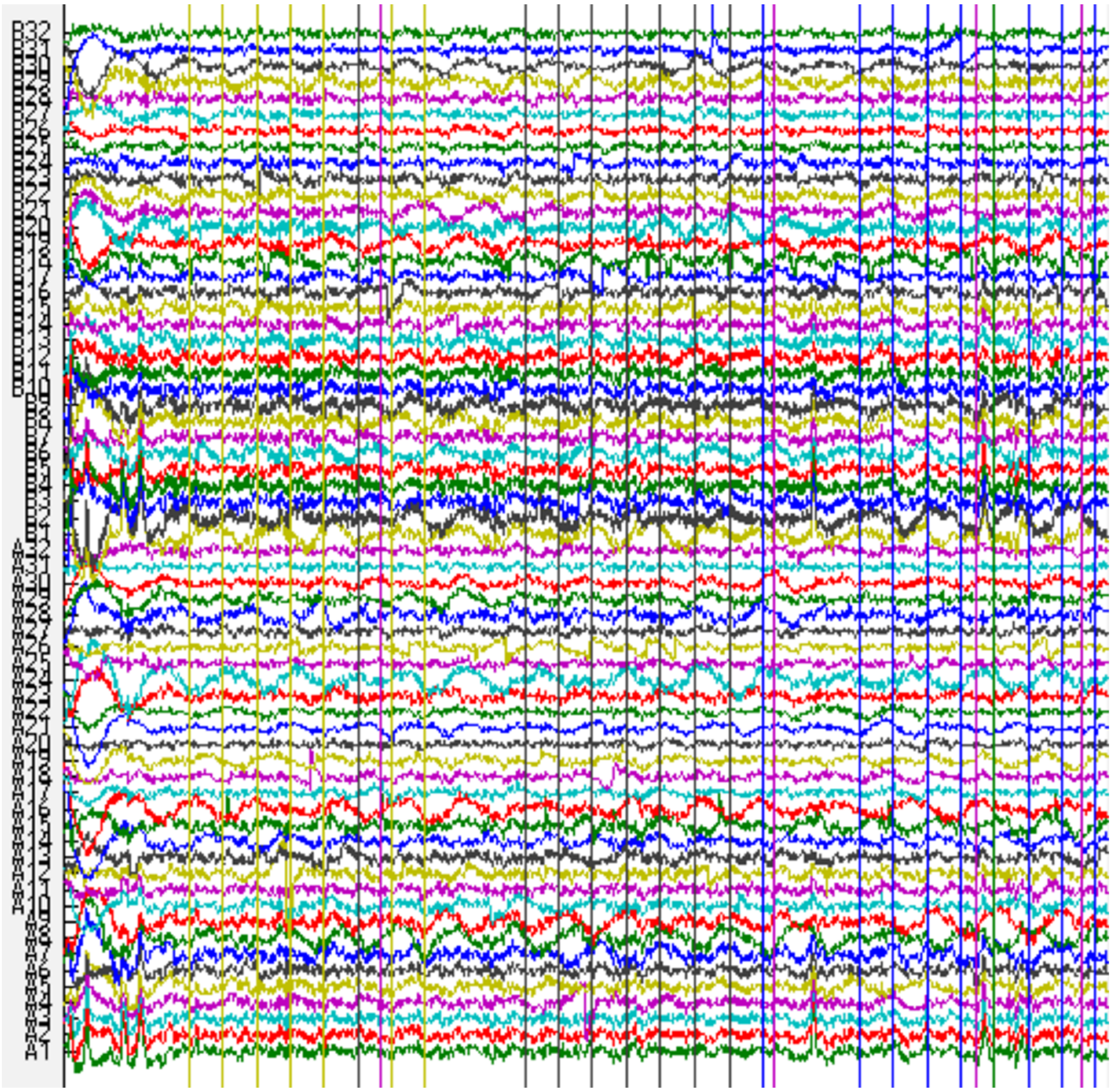

FIGURE 2.4: 64 channels EEG data were recorded by Biosemi

2.1.5.1 Imagery BCI Controlled Wheelchair

With the brain controlled wheelchair from RIKEN Brain-TOYOTA Collaboration Center (2008) shown on Figure 2.5 the user continuously controls the of the wheelchair by imagining left hand, right hand or both feet movements.

FIGURE 2.5: With the RIKEN's brain controlled wheelchair, the user continuously controls the velocity of the wheelchair by imagining left hand, right hand or both feet movements (Figure obtained permission)

2.1.6 Stimuli-driven BCI based on EEG

A Brain Computer Interface (BCI) is any system which can derive meaningful information directly from the user's brain activity in real time. The most important applications of the technology are mainly meant for the paralyzed people who are suffering from severe neuromuscular disorders. Most BCIs use information obtained from the user's EEG, though BCIs based on other brain imaging methods are possible. This section briefly describes several stimuli-driven BCIs.

2.1.6.1 Event Related Potential

Event-related potentials (ERPs) are very small voltages (micro-volt) evoked in the human brain in response to specific events or stimuli (such visual, auditory, tactile, olfactory) [21]. They are brain activity changes that are time locked to sensory, motor or cognitive events. The study of the brain in this way provides a noninvasive means of evaluating brain functioning in patients with cognitive diseases. ERP is a method of neuropsychiatric research which holds great promise for the future. Two special ERPs are introduced in this section.

Visual Evoked Potential (VEP):

- The EEG potential changes can be observed after visual stimulus. In the past decades, Adrian and Matthew are found this ERP response. The first nomenclature for occipital EEG components in 1961 are developed by Ciganek. After the Ciganek, Hirsch and colleagues recorded a VEP components on the occipital lobe, and they discovered amplitudes recorded along the calcarine fissure were the largest. An attempt to localize structures in the primary visual pathway was completed by Szikla and colleagues. Halliday and colleagues completed the first clinical investigations using VEP by recording delayed VEPs in a patient with retrobulbar neuritis in 1972. A wide variety of extensive research to improve procedures and theories has been conducted from the 1970s to today.

Auditory Evoked Potential (AEP):

- The EEG potential changes can be observed after auditory stimulus. Auditory evoked potentials (AEPs) are a subpart of event-related potentials

(ERP)s. ERPs are brain responses that are time-locked to some specific events or stimuli, such as a sensory stimulus, a mental event (such as recognition of a *target* stimulus), or the omission of a stimulus. For AEPs, the event is a sound, speech or music. AEPs (and ERPs) are very small electrical voltage potentials (micro-volt) originating from the brain recorded from the scalp in response to an auditory stimulus, such as different tones, speech, sounds, noise etc.

2.1.6.2 P300 Response

The $P300$ response is an event related potential (ERP) component elicited in the process of decision making. It is considered to be an endogenous potential, as its occurrence links not to the physical attributes of a stimulus, but to a person's reaction to it. More specifically, the $P300$ is thought to reflect processes involved in stimulus evaluation or categorization. It is usually elicited using the oddball paradigm, in which low-probability *target* items are mixed with high-probability *non-target* (or "standard") items.

When recorded by electroencephalography (EEG), as a positive deflection in voltage with a latency (delay between stimulus and response) of roughly 250 to 500 ms. The signal is typically measured most strongly by the electrodes covering the parietal lobe. The presence, magnitude, topography and timing of this signal are often used as metrics of cognitive function in decision making processes. While the neural substrates of this ERP component still remain hazy, the reproducibility and ubiquity of this signal makes it a common choice for psychological tests in both the clinic and laboratory.

In 1988, Farwell and Donchin developed the first $P300$ based BCI to select letters from a virtual keyboard. Items are presented on a 6 by 6 matrix; rows and columns are flashed in a random sequence, eliciting a $P300$ signal at 300 ms after the key the user wants to select has been flashed.

2.1.7 Review of stimuli-driven BCI

In this section, I introduce two representative stimuli-driven BCIs. I focus on the status of the $P300$-BCI first. This system has now been tested with several dozen

ALS patients and some have been using this approach for communication at a very extensive level. The detail deception shows in the next section.

2.1.7.1 Visual Keyboard Speller BCI

A $P300 - based$ spelling device, originally created by Farwell and Donchin in 1988 and improved by several lab groups since its inception [22], [23], has had considerable success in allowing these patients to communicate efficiently with others. The underlying principle of the $P300$-based spelling system is the "oddball paradigm", a term used to describe a specific event related potential (ERP) that follows presentation of rare stimuli amongst expected or predicted stimuli. This ERP typically occurs approximately 300 milliseconds after the presentation of an unexpected stimulus, thus allowing a stimulus to be discerned from others. Using this principle, characters are presented to the user of the $P300$-based spelling device in a 6x6 matrix on a black screen. The characters are then highlighted in a random, flashing sequence by row and by column, respectively (see Figure 2.6). Users are instructed to focus attention on only one character in the matrix, the "*target* character". This selective attention to one character is then evident due to the presence of the ERP 300 milliseconds after the character's corresponding row and column were highlighted. It is therefore possible to identify the targeted character; the character's location is at the intersection of the row and column numbers with the $P300$ ERP results. Noted successful methods of users for $P300$-based spelling devices include silently counting the number of times the attended 57 character was randomly highlighted or silently affirming that it had been highlighted at the moment the flash occurred. The $P300$-based spelling device allows the user communication via character selection and therefore enables him or her to type without a requirement for voluntary motion. Although $P300$-based technology is seen as a reliable method for communication, it has never been applied to enable the use of an internet browser. However, for individuals already accustomed to the $P300$-based spelling system, the transition to the internet could combine the familiarity of a system they can already operate with the freedom of the world wide web, opening many opportunities for those who are severely paralyzed.

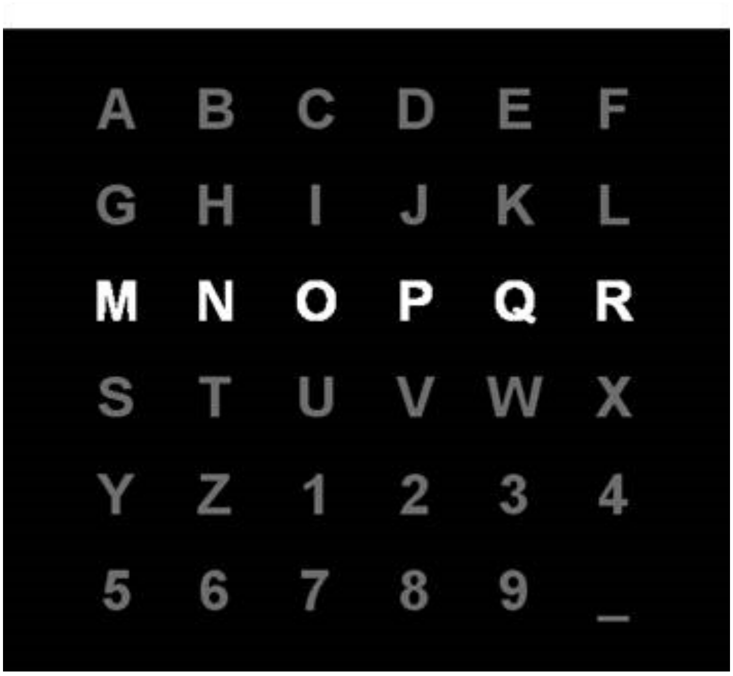

FIGURE 2.6: In the $P300$ speller by Farwell and Donchin, items are presented on a 6 by 6 matrix. Rows and columns are flashed in a random sequence, eliciting a $P300$ signal 300 ms after the key the user wants to select has been flashed. The figure source: http://www.bbci.de/competition/ii/

Chapter 3

Electrodes Position Optimization

In this chapter, I introduce a novel auditory BCI paradigm based on combined sound timbre and horizontal plane spatial locations as informative cues. The presented concept is based on responses to eight–directional audio stimuli with various tonal and environmental sound stimuli. The approach is based on a monitoring of brain electrical activity by means of the electroencephalogram (EEG). The previously developed by the authors spatial auditory stimulus is extended to varying in timbre sound stimuli which feature helps the subjects to attend to the *targets*. The main achievement discussed in Chapter 3 is an offline BCI analysis based on an optimization of electrode locations on the scalp and evoked response latency for further classification results improvement. The so developed new BCI paradigm is more user-friendly and it leads to better results comparing to previously utilized simple tonal or steady-state stimuli.

3.1 Introduction

A concept of a spatial auditory stimulus creates a very interesting possibility to *target* "the less crucial" auditory activity. I propose to utilize spatial audio stimuli design with a *target* application in a new BCI paradigms where users consciously direct their attention to different locations in surround sound environment with various tonal frequency stimuli [24], as depicted in Figure 3.1. Contemporary applications limit their scope to frontal surround sound loudspeakers [12], while our proposal includes also rear loudspeakers sound presentation allowing for eight commands BCI applications (full octagonal surround sound loudspeakers setup).

In the approach first proposed in [25] it was shown that responses in a spatial tonal stimuli within the 7.1 channels surround sound system (subjects were positioned in the middle of the loudspeakers systems and requested to direct attention to single direction loudspeakers) were distinguishable in EEG for *targets* and *non-targets* interfacing commands. The *target* and *non-target* direction sequences were presented randomly. The current proposal extends the design to fully octagonal loudspeakers setup with stimuli direction sequences presented also randomly.

Within this framework, the subjects are asked to focus their attention to a direction of the tonal or environmental sound. The EEG responses are recorded with an EEG amplifier. Additionally vertical and horizontal eye-movements are recorded in order to have a reference signal indicting potential muscle activity used later in artifacts removal algorithm. In the presented study I decided first to process only "artifact-free-data" (eye blinks, facial muscle and head movements, etc. trials were discarded) in order to validate the proposed spatial auditory stimuli paradigm. To reduce computational complexity of the interfacing approach I propose channel and event-related-potentials (ERP) response samples selection in order to optimize classification results. This approach allows for EEG channels selection optimization and ERP regions-of-interest (ROI) for *target* and nontarget stimuli optimization by adaptively finding only the components carrying brain activities maximizing the contrasts between responses when the subjects attend or ignore the spatial stimuli. The so obtained brain activity spatial patterns clearly follow the expectations of stronger activities in parietal and temporal cortical areas, which are known for spatial stimuli processing [26]. The chapter is organized as follows. In the next section the experimental paradigm is described together with EEG recording introduction and EEG preprocessing steps. Next the channel selection and ERP response period optimization procedures for each subject are described. Finally classification results and discussion conclude the chapter.

3.1.1 Methods

The experiments to validate the proposed spatial auditory BCI paradigm were conducted in a Laboratory for Advanced Brain Signal Processing, RIKEN Brain Science Institute in Wako-shi, Saitama, Japan with agreement of the institute's ethical committee guidelines. All experimental procedures and this study targets were explained to the subjects who agreed to participate voluntarily by signing

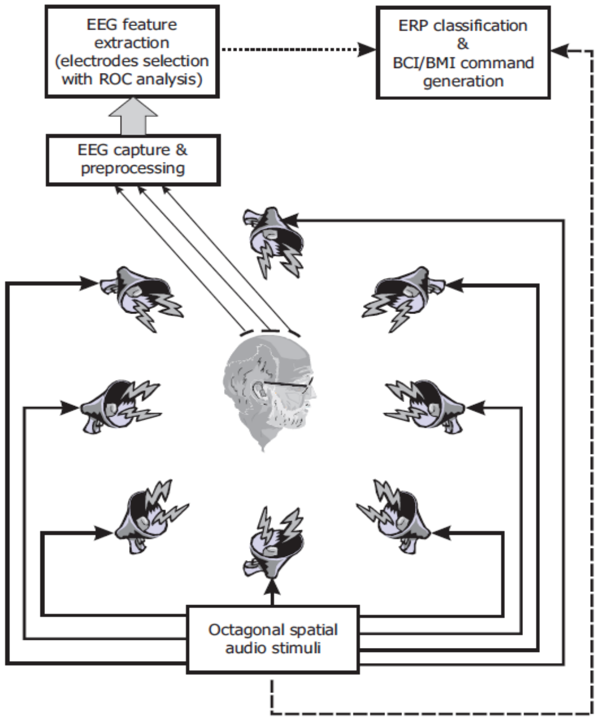

FIGURE 3.1: Spatial auditory BCI paradigm concept with octagonal loudspeaker arrangement.

consent forms. EEG signals were recorded with 64 channels active electrodes EEG caps with BIOSEMI ActiveTwo system (see section EEG recording system). The stimuli sounds were played through loudspeakers positioned in octagonal setting around a head of a subject. All experiments were conducted in a silent and low reverberation room (see figure 3.2)in order to limit an interference of "environmental noise" in this preliminary study in order to validate usability of all front- and rear-head sound usability (in contrast to a study conducted in a noisy environment which failed to utilize back loudspeakers for BCI application [12] - only psychoacoustic experiment for rear-head loudspeakers succeeded there).

A diagram of stimuli and further described in this section steps of EEG processing, electrodes positions optimization, feature extraction and final classification are depicted in Figure 3.1.

3.1.1.1 EEG Recording System

As I mentioned above, the BIOSEMI ActiveTwo system was used for our EEG experiment. The BioSemi ActiveTwo measurement system is designed to measure

FIGURE 3.2: Spatial auditory BCI open acoustic filed.

potential differences on the human body surface. The system is successfully used to record signals originating from the brain (electroencephalography, EEG) and the muscles (electromyography, EMG) for research purposes. The ActiveTwo system can be adapted to these different applications by using different versions of the (active) electrodes. The BioSemi ActiveTwo is designed to digitize the signals from 8 up to 256 active electrodes and other sensors.

3.1.1.2 Stimulus

Two types of auditory stimuli were presented spatially to the subjects in the octagonal speaker system through only single speaker at a time. The first stimuli was composed of a 400 ms long (the longer stimuli comparing to results from [12] was chosen to create more spatial localization realistic situation) and 440Hz sinusoidal tone with 10 ms linear raise and decay to avoid "click-effect". The low frequency tone was chosen to evaluate feasibility of an inter-aural-time-delay (ITD) brain auditory localization principle only [27]. The second stimuli was a sound of a car horn recorded on a noisy street and presented also in a form of 400 ms long waveforms with 10 ms linear raise and decay periods. This second stimuli represented a "broadband" acoustic waveform targeting both brain auditory mechanisms of inter-aural-level-difference (ILD) and ITD. The sound directions were presented in random order to avoid habituation effects.

3.1.1.3 Spatial Hearing

Localization of sounds in space is one of the processes that human brain can do without mental effort [28]. Several studies show the ability of human to distinguish sounds in space [29], [30], [31]. Two studies showed that when people focus on a direction, their attentional resources appear to be distributed in a gradient, with decreasing alterness when moving away from the attended direction [30], [31]. I propose to utilize a concept of a horizontal two dimensional sounds spatial sources arrangement of stimuli paradigm in order to evoke $P300$ response discriminating subjec's intentionally attended *target* and *non-target* directions. I test the hypothesis whether the evoked in EEG ERP signals are feasible for a future multi-command (eight commands in the proposed case) BCI application similarly as it was developed in visual domain [32]. In order to present the spatial auditory stimulus, I used the software Max/MSP (see Figure 3.3, 3.4). Max is a visual programming language for music and multimedia developed and maintained by San Francisco-based software company Cycling '74.

3.1.1.4 EEG Experiment Description

EEG recording experiment for offline BCI paradigm testing were conducted with six healthy subjects (five males and one female; age range $20 - 50$ years). The subjects were instructed to sit in a comfortable chair in center of eight octagonally positioned loudspeakers. The elevation of the loudspeakers was fixed to the subjects' ear level. Instructions which *target* direction to attend was given visually on a display located in front of them. A visual fixation cross was also presented on that display to avoid unnecessary eye movements.

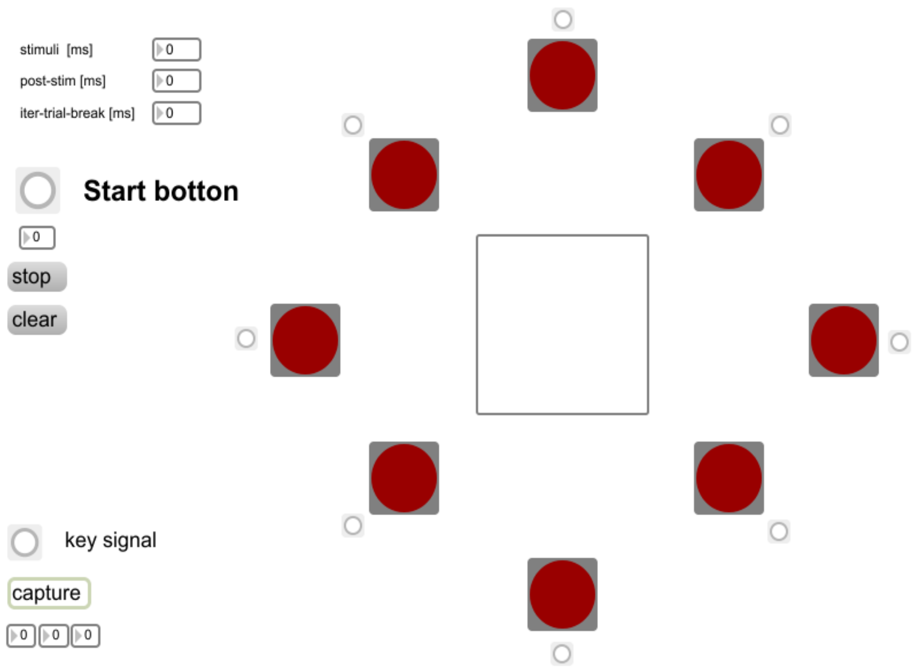

FIGURE 3.3: User training interface for spatial auditory BCI experiment.

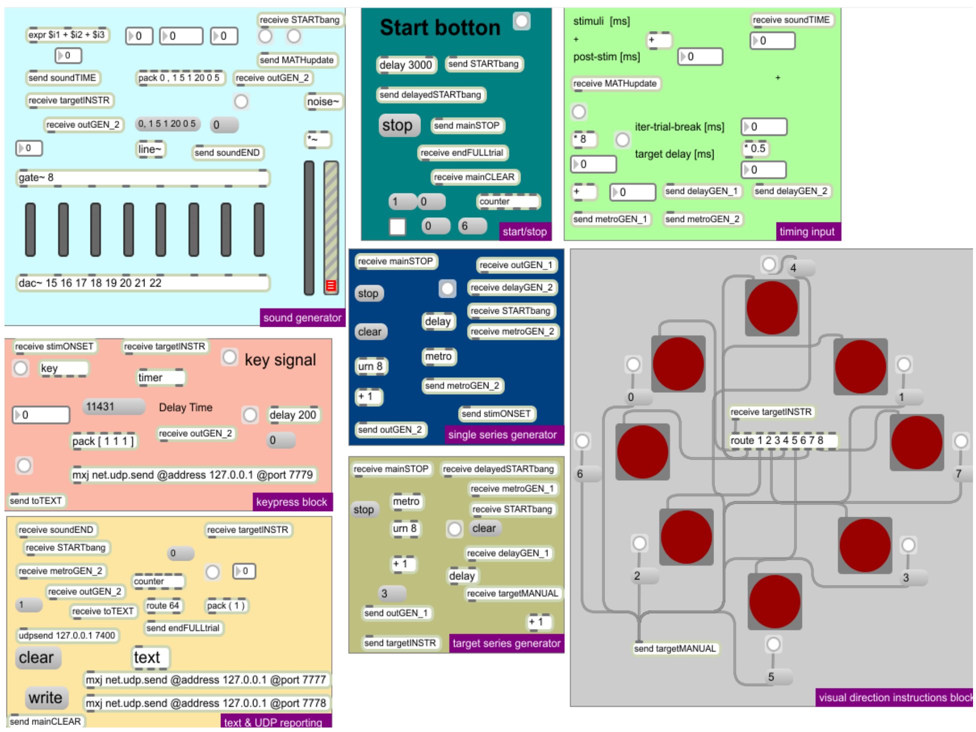

FIGURE 3.4: MAX/MSP path managing the spatial BCI Experiment.

3.1.2 Analysis

The EEG analysis leading to final eight-directions-spatial auditory classification for *target* and *non-target* locations is composed of three steps as follows:

- a EEG signal preprocessing (filtering and artifact removal);
- an informative electrodes selection and ERP's ROI optimization for further classification outcome maximization;
- a final classification of evoked responses within each of ten chosen "best channels";

3.1.2.1 EEG Preprocessing

The recorded raw EEG 64 channel (see figure 3.5) signals with BIOSEMI ActiveTwo system have to be first referenced by removing mean values of all channels. Next a notch filter is applied at 50 Hz center frequency to remove power line noise interference. Two Butterworth 5th order low-pass and highpass filters are applied next with cutoff frequencies at 0.5 Hz and 25Hz respectively to remove low frequency and DC-shift interferences. The low–pass filtering removes possible muscle frequency artifacts. Next the signals are segmented creating event related epochs starting at 0 ms of the stimulus onset and ending at 500 ms after it (see Figure 3.6). In a next step eye movement artifacts rejection is carried out. Spatial stimuli are known to cause uncontrolled eye movements which in the current approach are removed with a threshold value set at 80μV (signal voltage above EEG activity level). The EEG conversion from the original BIOSEMI BDF format and the above preprocessing steps were conducted within SPM8 package [33]. The rejected epochs are not processed further since in current approach an emphasis is put forward on the spatial paradigm validation. An example with artifact cleaned and averaged epochs separately for frontal and rear loudspeakers is presented in Figure 3.6. This confirms a feasibility to utilize all frontal and rear channels, since in averaged ERPs those responses are clearly easy to distinguish in $300 - -500$ ms time range.

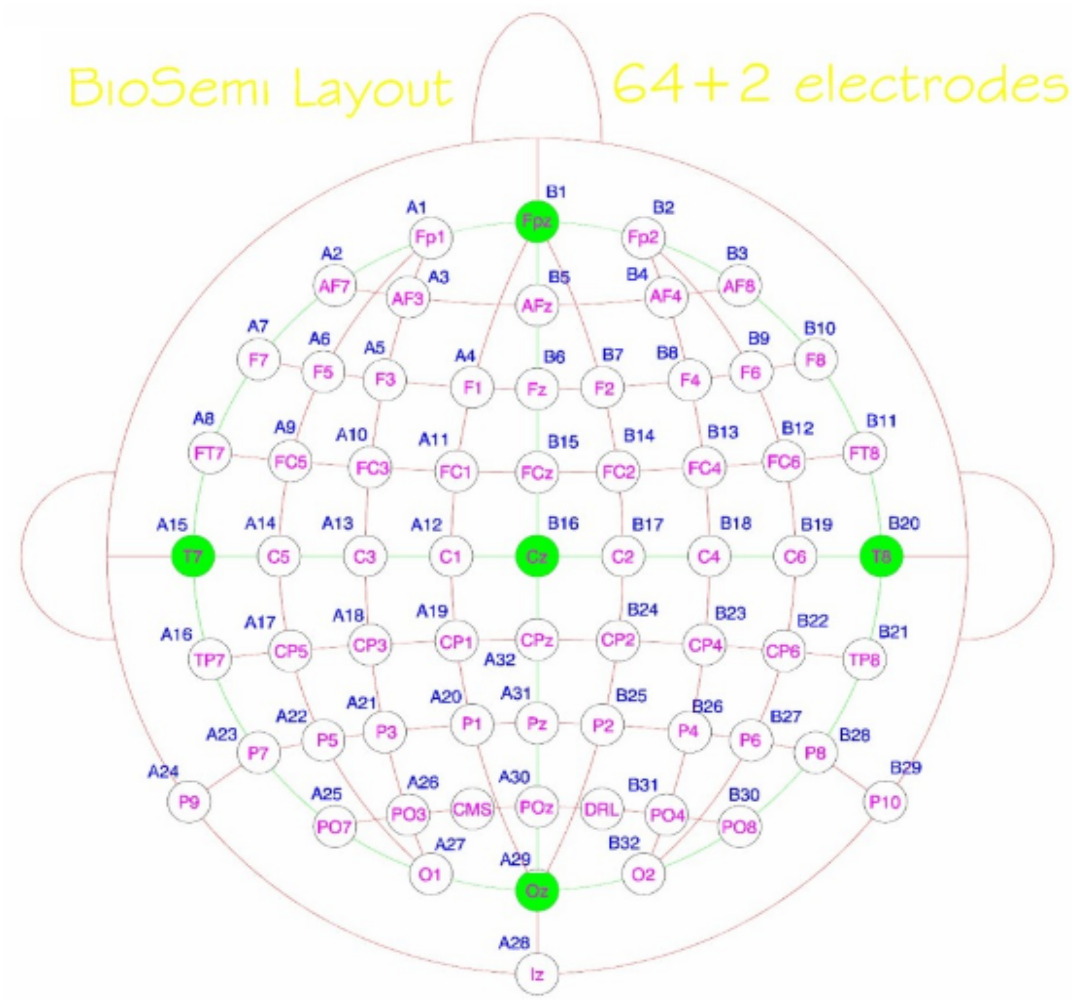

FIGURE 3.5: BIOSEMI system with 64 electrodes.

3.1.2.2 Electrodes and ERP Features Selection for Classification

In order to find ten electrodes from the original 64 channels recordings which for each subject would discriminate the ERP responses for *target* and *non-target* responses, I propose to test the two measures. The first one is based on a classical linear-discriminant-analysis (LDA) [34] classification applied to all channels separately. The best classification results set of ten channels from 64 available in a training set would be later applied to test sets. The second proposed measure is based on a receiver operating characteristic (ROC) [35] applied to quantify the separability of two single-trial response distributions for each sample point of

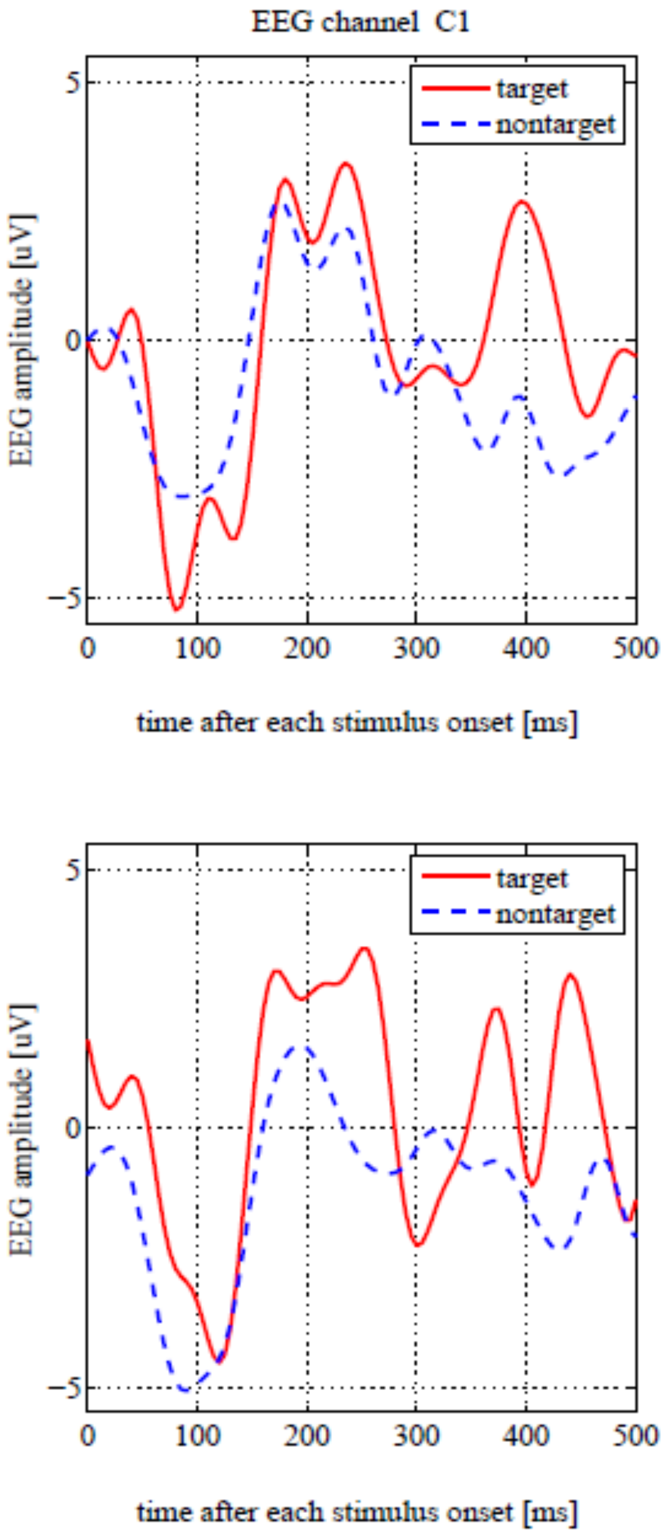

FIGURE 3.6: Results of ERP $P300$ response for frontal (upper panel) and rear (lower panel) speakers confirming the feasibility of the proposed approach. The zero stands for stimuli onsets. The blue/dashed lines depict *non-target* (no $P300$ response after 300 ms) responses and red/solid traces visualize attended spatial *targets* (obvious positive EEG response deflections after in 300 – 500 ms range).

ERPs (see Figure 3.6 with averaged ERP responses for *targets* and *non-targets* suggesting that ROI there shall be chosen from around $350-500$ ms). While LDA is a standard classification method the ROC related measures are usually used to evaluate the performance of classifiers, they can also quantify the discriminability of feature distributions leading to classifiers optimization. The ROC curve derived from perfectly mixed distributions is the diagonal line (a no–discrimination line as presented approximately in Figure 3.7 for not chosen channel). Analysis conducted with [36], [37]. The numbers along the major diagonal of ROC graph represent the correct decisions made, and the numbers off this diagonal represent the errors of the confusion between the classes. The sensitivity (also called a true positive rate) of a classifier is calculated as:

$$sensitivity = \frac{PCC}{TP}, \tag{3.1}$$

and the classifier specificity is as:

$$specificity = \frac{TN}{TP+TN}, \tag{3.2}$$

where PCC stands for positives correctly classified; TP for total positives; TN for true negatives; and FP for false positives respectively. The results of ROC analysis for the chosen and discarded EEG channels are presented in Figure 3.7.

In order to choose channels with EEG ERP features leading to best classification results I utilized a separability index which is calculated as an area under the curve (AUC) between the ROC curve and the no-discrimination line (diagonal) multiplied by two to relate it to the Wilcoxon test of ranks or the Gini coefficient [35]. I decided to choose ten channels scoring with highest AUC for each subject and within each experimental paradigm (440 Hz tone and car horn in the current study). Additionally within each of the chosen channels the best discriminable two areas were chosen taking only 50 ms regions around AUC maxima derived from $0-200$ ms and $200-500$ ms regions. Those two vectors of single-trial ERP subregions formed features used in subsequent LDA classification within each channel.

Finally the EEG ERP responses were classified into *targets* and *non-targets* using two approaches to validate the proposed electrodes and ROI estimation with ROC together with classical LDA applied to all electrodes and the $P300$ response

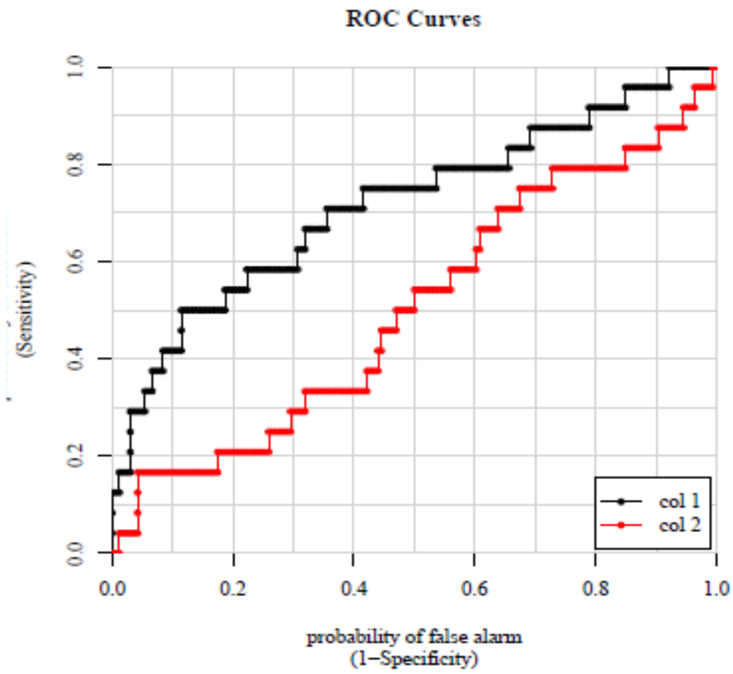

FIGURE 3.7: ROC analysis results of a good channel candidate (col1: black-/upper trace) discriminating and a "chance level" one for not channel selection (col2: red/lower trace).

region. For classical LDA the results are visualized in Figure 3.8, where classification outcomes are shown for all electrodes together with the ten best electrode candidates. For the proposed approach of ROC based channel and ERP ROI selection the best electrode candidates are visualized in Figure 3.9, where the ten best electrodes are indicated in red for each subject and condition. The detailed results are discussed in the next section.

3.1.3 Results

The results of the proposed approach to compare *target* and *non-target* evoked potentials have been summarized in Figures 3.10 and 3.11. In Figure 3.10 it has been shown that the proposed approach to identify the ten best electrodes and ERP response ROI based on ROC analysis had allowed for a gain of classification results ranging from an increase of 8% boost at the best for subject #1 and tonal stimuli of 440 Hz comparing to classical application of LDA analysis to $P300$ response area for all electrodes and the whole ERP region. In case of a car horn stimuli, the best classification increase has been obtained also at the level of 5%

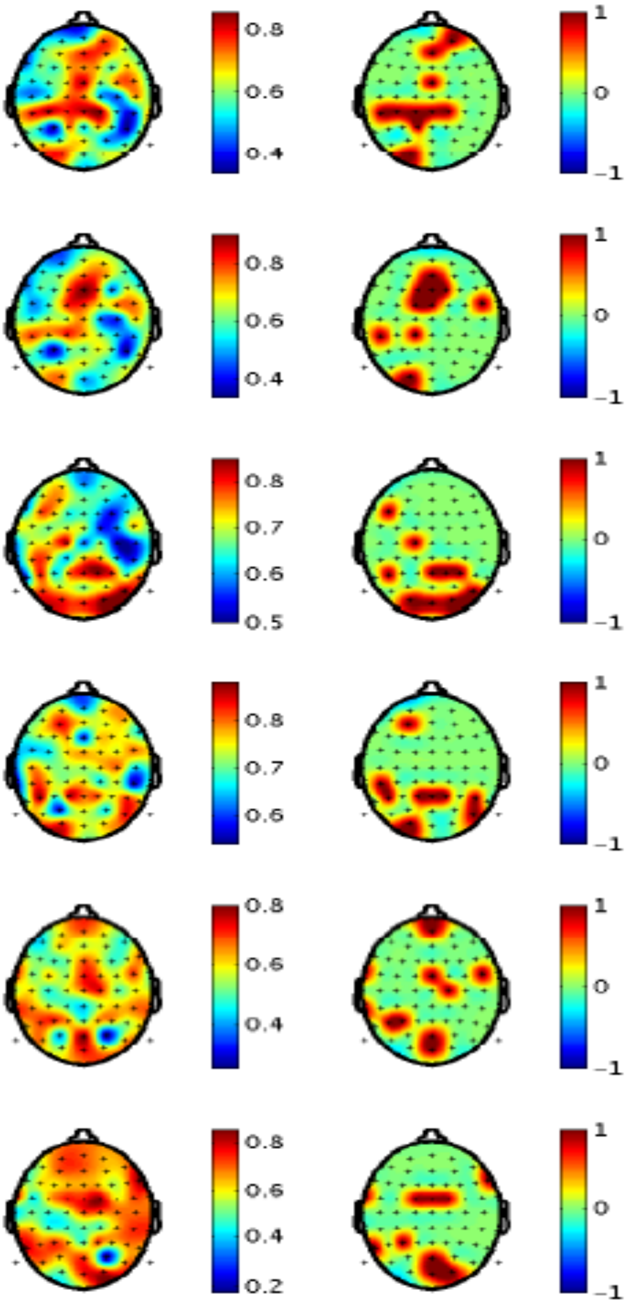

FIGURE 3.8: Results of classical LDA application to binary classification of a spatial tonal 440 Hz stimuli applied to all electrodes for a single subject in six crossvalidation trials are visualized in the left column while the best resulting electrodes are presented in the right column for each subject. All graphs have the same scaling..

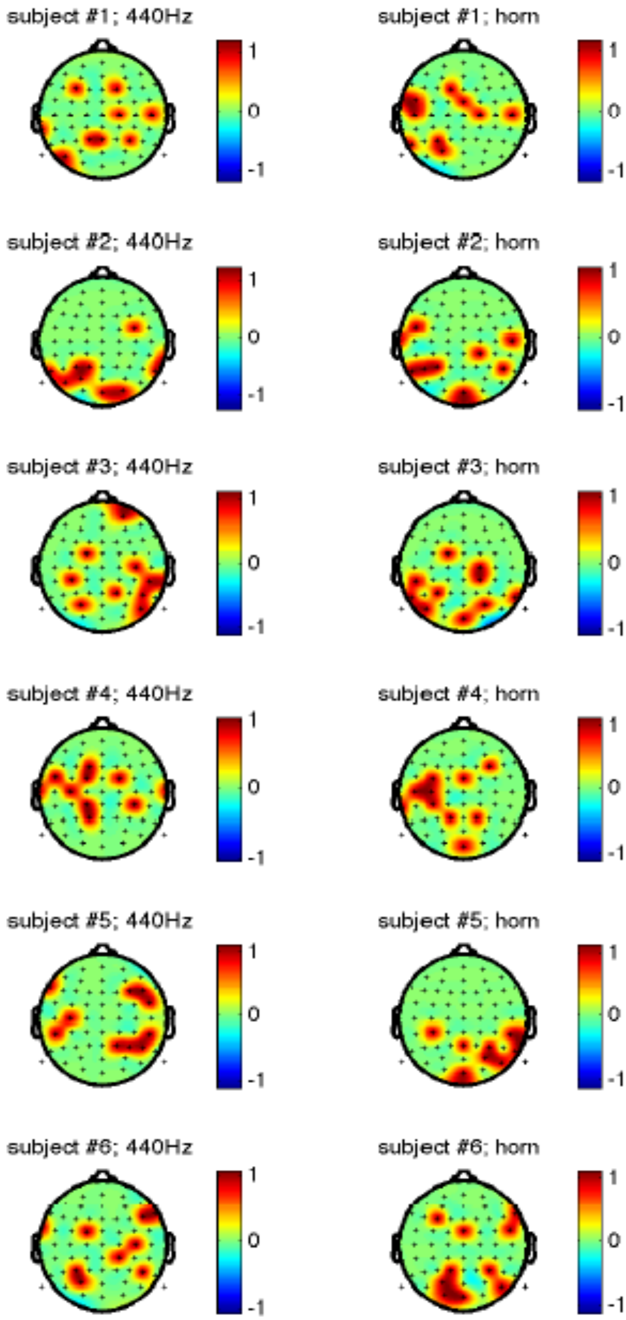

FIGURE 3.9: Results of the proposed ROC analysis based channel selection for all six subjects and two stimuli cases revealing the temporal and parietal scalp regions as best candidates for spatial stimuli $P300$ responses identification.

for the same subject. Figure 3.11 presents a comparison of *target* vs. *non-target* classification results for frontal and rear loudspeaker sound directions confirming the hypothesis of a possibility to utilize those direction modalities despite of the known in psychoacoustics *front-back-confusion* effect. A variability of the results for various subjects and conditions calls still for further research in this area which our group will continue.

3.1.4 Conclusions

In the chapter it has been shown that in contrary to the contemporary results with the spatial auditory BCI paradigms, which fail to utilize rear-head loudspeakers, it is possible to achieve good results for a fully surround sound octagonal loudspeakers setup. The developed by our approach to select the optimal ten channels and ERP ROI intervals resulted with very good classification results for all eight sound stimuli directions. This has been achieved for two types of 400 ms long acoustical stimuli targeting ITD and ILD auditory spatial localization

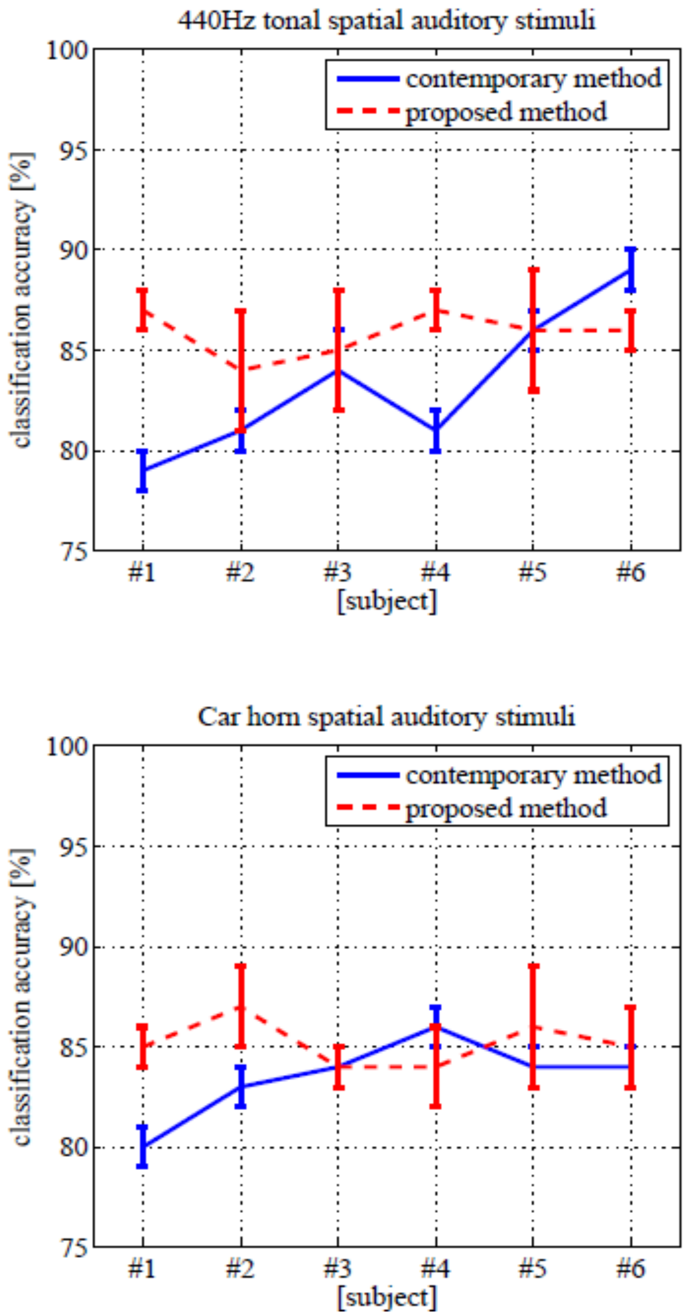

FIGURE 3.10: Percentage of correctly classified spatial audio BCI $P300$ responses with LDA classifiers derived from three experimental sessions for each subject (blue/solid lines) and based on ten best electrode results (red/dotted lines). Chance level is 50%.

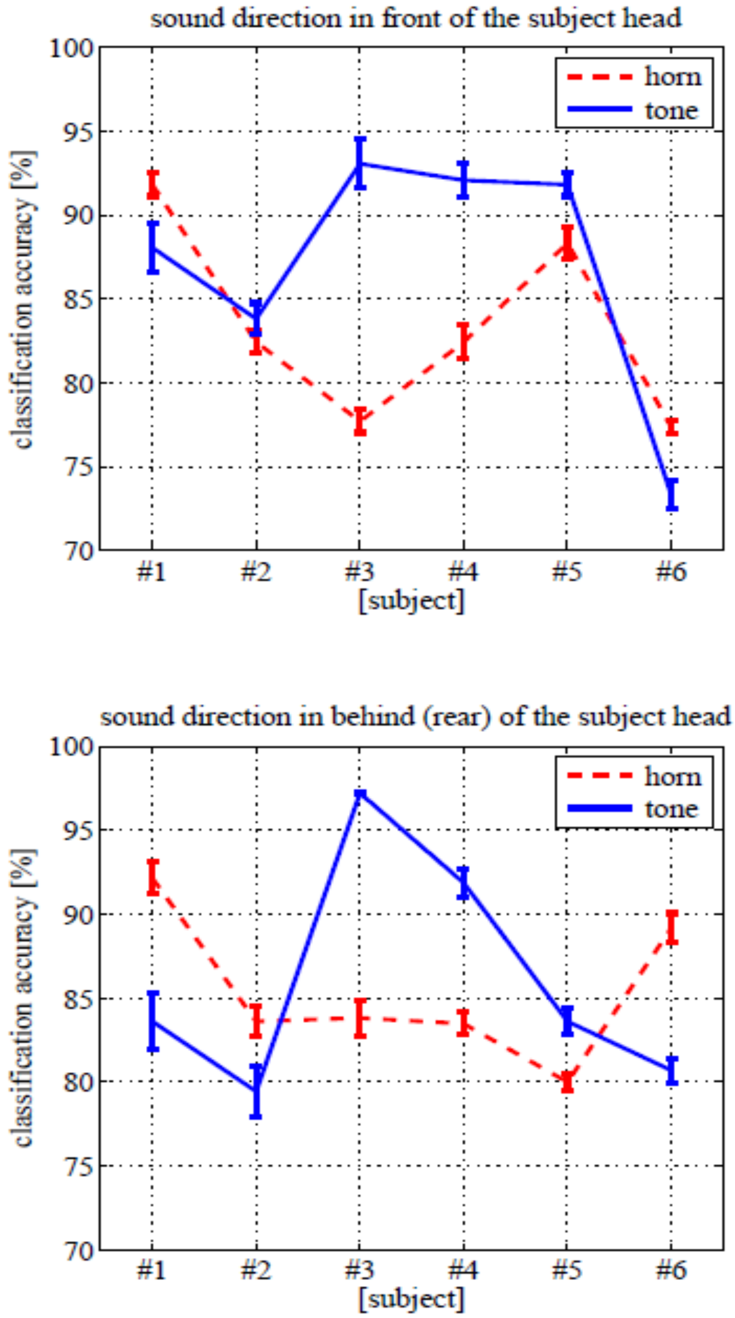

FIGURE 3.11: Comparison of classification responses to frontal (top panel) and rear/back (bottom panel) loudspeakers stimuli directions for six subjects and two stimuli conditions. The results confirm only slight subjects' preferences to frontal stimuli directions except of subject #6 who had better results for rear sound directions..

Chapter 4

$P300$ Optimization for Spatial aBCI

In this chapter, I propose a novel method for the extraction of discriminative features in electroencephalography (EEG) evoked potential latency. Based on our offline results, I present evidence indicating that a full surround (eight–directions) sound auditory BCI paradigm has potential for an online application. The auditory spatial BCI concept is based on chapter 3, which employs a loudspeaker array in an octagonal horizontal plane. The stimuli presented to the subjects vary in frequency and timbre. To capture brain responses, I utilize an eight-channel EEG system. I propose a methodology for finding and optimizing evoked response latencies in the $P300$ range in order later to classify them correctly and to elucidate the subject's chosen *targets* or ignored *non-targets*. Finally, I identify the subjects' intended commands with a naive Bayesian classifier for sorting the final responses. The results obtained with ten subjects in offline BCI experiments support our research hypothesis by providing higher classification results and an improved information transfer rate compared with state–of–the–art solutions.

4.0.5 Introduction

In chapter 3, I discussed EEG electrodes selection, event-related potential (ERP) features optimization and linear discriminative analysis classification [38]. These earlier results were unsatisfactory as regards a successful online aBCI system application.

In this chapter, I discuss a auditory BCI paradigm based on the full surround sound horizontal stimuli as an informative cue with the support of the $P300$ component at a latency around and after the 300 ms, evoked by the expected/instructed *targets* [14]. Our hypothesis is that a significant ERP response will be found when subjects attend to the *target* direction and ignore the *non-targets*. To find the significant differences, I propose analyzing the response statistically to identify only those ERP latencies that contribute to the classification enhancement, in contrast to state-of-the-art approaches [1], in which the whole response is taken as a feature for subsequent classification.

The hypothesis of our research is that the horizontal full surround sound aBCI paradigm could be improved with the careful selection of ERP discriminative features that allows the use of the rear-to-the-head sound directions. For this purpose, I introduce a statistical response analysis, which ultimately leads to the final improvement in the information transfer rate (ITR).

The chapter is organized as follows. In the next section, the experimental paradigm is explained together with the EEG preprocessing steps. Then, I discuss EEG feature selection using the method of statistical analysis of ERP responses. Finally, I present classification results obtained with a Gaussian Naive Bayesian Classifier (GNBC), which leads to an improvement in the ITR scores. The final section summarizes the chapter.

4.0.6 Method

Within the framework of the proposed novel aBCI paradigm, the subjects were asked to attend to and count *targets* while ignoring *non-targets*, as in the classical oddball paradigm [1], [8] , [9],[14]. A *target* direction instruction regarding which direction should be attended to in each trial was displayed visually on a computer display located in front of the subject. First, I conducted psychophysical experiments to check possible preferred directions of the subjects by comparing response time delays. Next I conducted EEG recording experiments in an offline BCI setting. The EEG signals were recorded with a g.MOBILab+ EEG amplifier by g.tec. I used novel dry EEG electrodes g.SAHARA to further improve the subjects' comfort, since these do not require conductive gel. The reference and ground electrodes were attached behind the left and right ears, respectively. To reduce unnecessary noise and to prevent degradation of the EEG signal quality as a result of EMG noise related to muscular movement in the ERP responses, the subjects were asked to minimize the blinking of their eyes, and facial and body movements during the experiments.

EEG experiments designed to validate the proposed spatial aBCI paradigm utilizing the $P300$ latency were conducted in the Multimedia Lab at the Life Science Center of TARA, University of Tsukuba, Japan. All the experimental procedures and study *targets* were explained to the subjects, who agreed to participate voluntarily. The experiments were conducted in agreement with the WMA Declaration of Helsinki?Ethical Principles for Medical Research Involving Human Subjects. All the experiments were conducted in a silent and low reverberation room in order to limit any interference from environmental acoustic noise.

The auditory stimuli were presented through eight loudspeakers in an octagonal setting, as depicted in the upper part of the Figure 4.1. The eight sound stimuli directions proved to be optimal from the points of view of aBCI and human subject spatial auditory performance [12].

Two short white and pink noise stimuli bursts were used as depicted in the lower part of the Figure 4.1 and described in the following section.

4.1 Psychophysical Experiment

In the psychophysical experiment, only the behavioral responses (button presses after the instructed and perceived *target* stimuli) were recorded. Different response time delays would suggest changing cognitive loads and task difficulties in function of the various spatial directions. The subjects were requested to press a button immediately after an instructed *target* direction was presented. The response delays in respect of auditory stimuli onsets were recorded and further analyzed in order to compare them with various spatial directions. The results of the psychophysical experiment are presented in Tables 4.1 and 4.2. As a result of the tests conducted, I conclude that all the eight spatial sound stimuli locations had the same (differences among means statistically non-significant when compared with pairwise t tests) values for all the tested octagonal stimulus spatial directions for white and pink noise, which also confirms psychophysical experiments reported in Schreuder et al. [12].

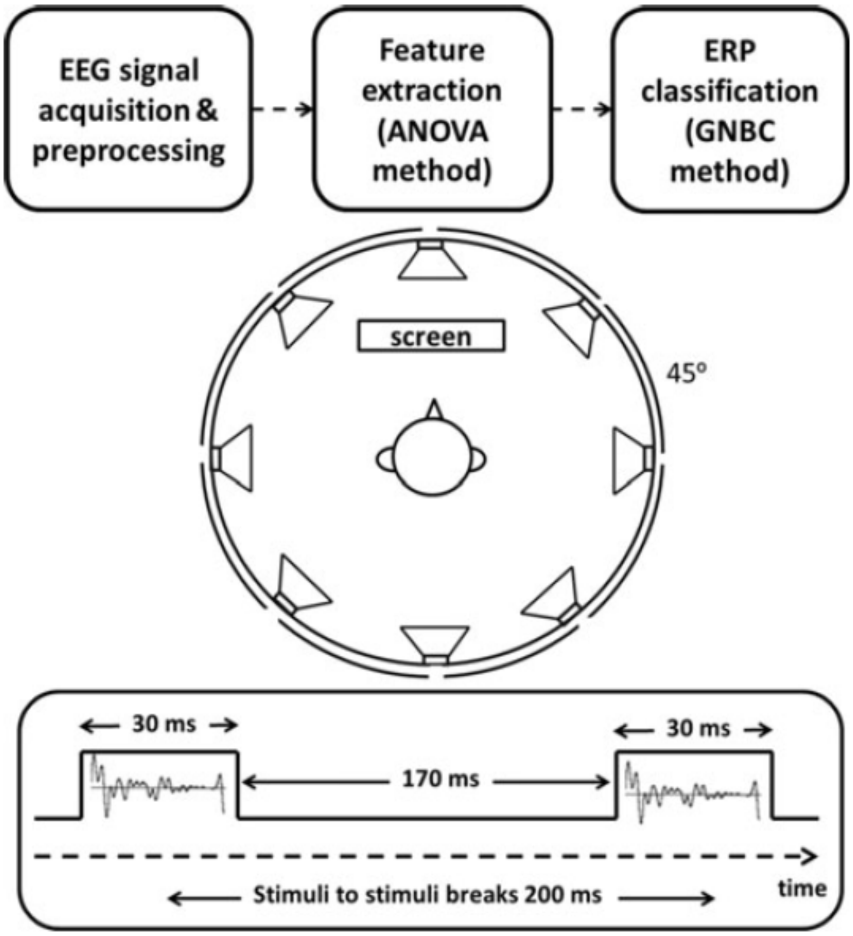

Figure 4.1: Spatial auditory BCI paradigm concept with eight loudspeakers in the upper part of the figure. The lower graph visualizes the stimulus presentation concept in the time domain. Each stimulus is presented for 30 ms with 170-ms silent breaks, so the ISI is set to 200 ms.

4.1.1 The Offline aBCI Experiment Protocol

The experimental hypothesis was that I would be able to distinguish from the ERP shape (mainly based on the *P*300 response latencies) which direction the subject attended to in the spatial auditory paradigm experiment. To test the hypothesis, I conducted a series of EEG recording experiments in the offline BCI mode (with no instant feedback or classification results given to the subjects [1]). EEG recording experiments were conducted with the ten healthy subjects (eight males; two females; age range from 23 to 42 years, mean 25.8, SD 6.34). The subjects were requested to sit in a comfortable chair in the center of eight octagonally positioned loudspeakers, and the dry EEG electrodes were positioned on the scalp. The elevation of the loudspeakers was fixed at the subject's ear level in order to create a horizontal spatial plane defined by the eight loudspeakers (see Figure 4.1). The volume of the sound was set to 72 dB.

The sound stimuli were presented in random order and one at a time from a single loudspeaker (a single trial consisted of a delivery of a single *target* and seven *non-targets*). I employed two broadband noise stimuli types that allowed us to utilize the two spatial localization mechanisms of the human auditory pathway, the internal time delay (ITD) and the internal level difference (ILD) [27]. The white and pink noise stimuli both had 30 ms lengths with 5 ms linear attack and sustain intervals. For each subject and each stimulus, I performed eight sessions (altogether 64 *targets* and 448 *non-targets* were presented). Each subject was requested to focus on the instructed *target* direction which was presented on a computer display. The subject ignored the other nontarget directions. Each subject was also requested to control her/his eye movements to decrease the unnecessary EMG noise during the experiments. Before each experiment, the subject was allowed a short practice session to get familiar with the spatial auditory stimulus conditions.

4.1.2 EEG Acquisition

The EEG signals were recorded by the g.MOBILab+ bio–amplifier with eight dry g.SAHARA electrodes. The EEG recording system captured the neurophysiological signals in a frequency range of 0.1–40.0 Hz. The following eight EEG electrode positions were chosen $P3$, $P4$, $P5$, $P6$, Cz, CPz, Pz, and POz (see figure 4.2), as in the 10/10 system [39]. The eight EEG channels were sampled with 256 Hz frequency and stored using a custom application programmed in MATLAB and Simulink environments.

4.1.3 EEG Response Analysis

The analysis of EEG ERP responses, leading to the final eight-direction spatial auditory classification for *target* and *non-target* locations, was composed of the following three steps:

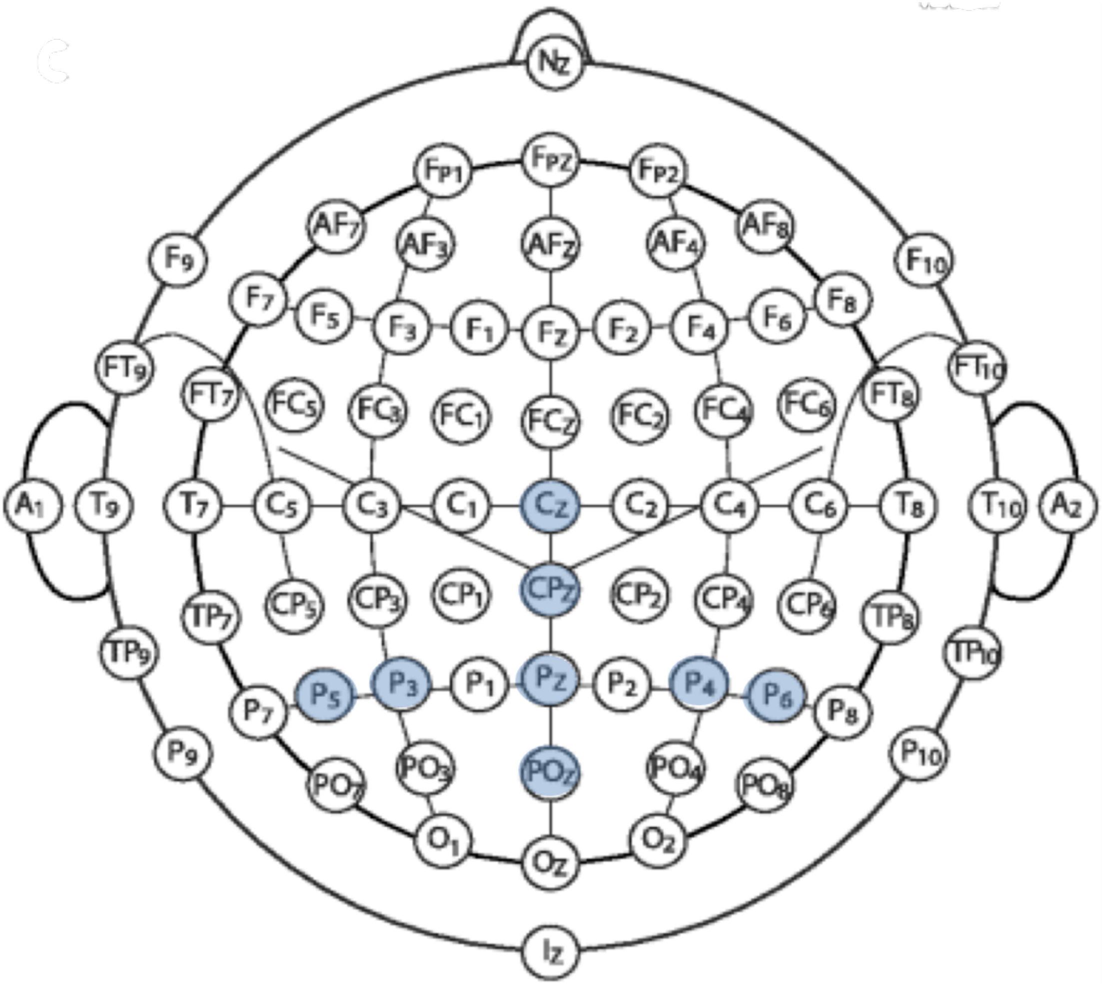

FIGURE 4.2: Eight electrodes on the scalp location (see the blue shadow)

- EEG signals preprocessing: band-pass filtering, epoch egmentation, and artifact rejection;
- Discriminative feature extraction using the analysis of variance (ANOVA) method;
- The final classification of evoked response using the GNBC.

I describe the above steps in detail in the following sections.

4.1.3.1 EEG Preprocessing

First, I filtered digitally the signals with the two fifthorder Butterworth high- and low-pass filters, which were applied with cut-off frequencies at 0.5 and 25 Hz. The lowpass filtering removed possible muscle-activity-related artifacts. The high-pass filtering removed the direct current-related drifts of the EEG signals, as well as slow eye movement artifacts.

Next, the EEG signals were segmented creating the ERP-related epochs. Each epoch started 100 ms before stimulus onset and it ended after 700 ms. I used the 100 ms prestimuli onset interval as the baseline (see Figure 4.3). In the next step, the rejection of eye movement artifacts was carried out. Auditory spatial stimuli are known to cause uncontrolled eye movements in subjects [40], which in the current approach were removed with a threshold value set at the 80μV (signal amplitude level above the usual EEG activity). The rejected epochs were not further processed, since in the current approach, the emphasis was on the spatial paradigm validation. In the following sections, feature extraction and ERP classification results are introduced.

4.1.3.2 ERP Feature Extraction Using ANOVA of the ERP Latencies

The aim here was to optimize the EEG response domain (mainly $P300$ response), which would provide a better separability for further classification. In order to do this, I conducted ANOVA of the two-class single-trial ERP distributions (*target* vs. *non-target* responses) in the spatial auditory experimental setting. The ERP response distributions passed "normality tests" and were comparable to more flexible methods such as the area under the curve analysis, yet the proposed ANOVA yielded the best results in our case. The majority of spatial aBCI applications aim at the $P300$ response latency [38], [41], [42], [43]. The example in Figure 4.1 shows the averaged ERP responses to *targets* and *non-targets* (note the latencies range $300 - 600$ ms). Next, the ANOVA method was applied to compare the differences of response distributions in single trials for each sample point of the collected ERPs. As a result, I were able to extract discriminative information leading to later classification optimization. The results of the above analysis are depicted in Figure 4.4 and 4.5. The bottom panels in the above figures visualize the ANOVA's p values for eight electrodes separately in each row using a

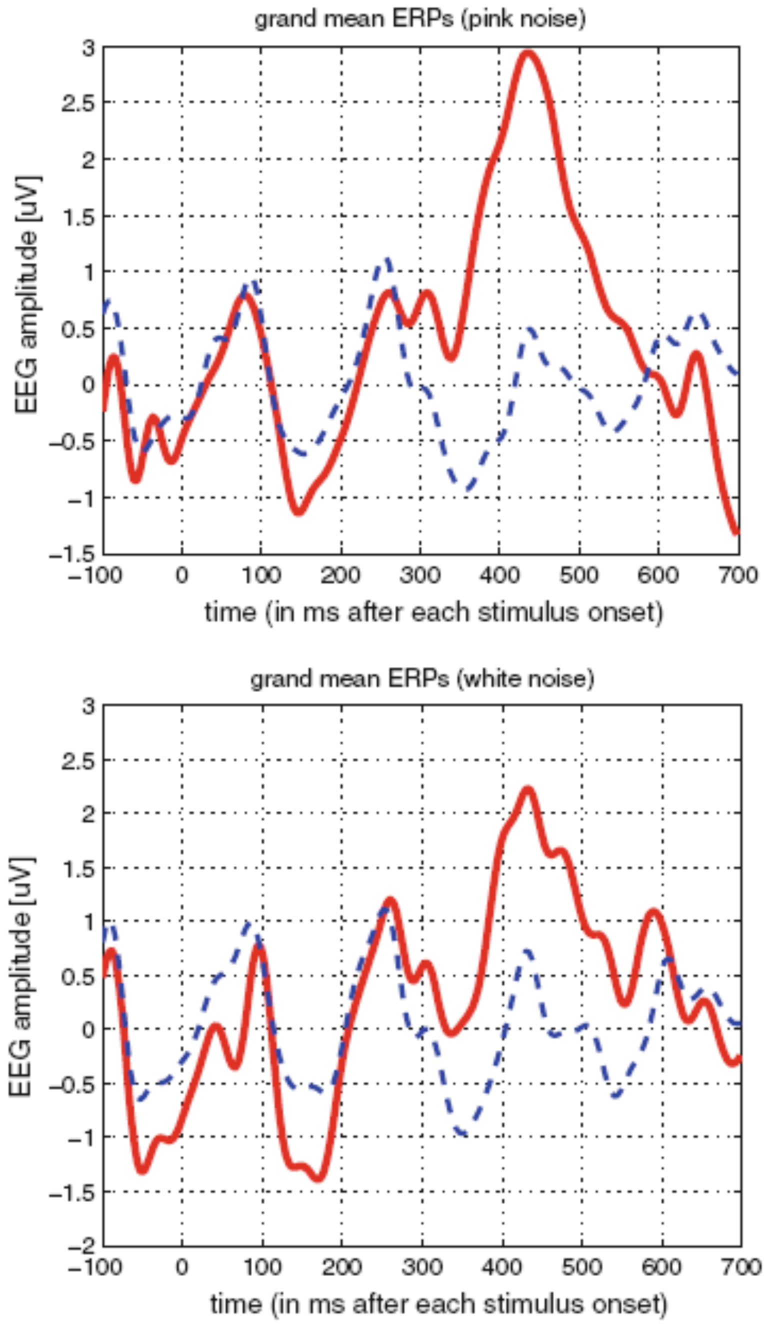

FIGURE 4.3: Spatial auditory BCI paradigm concept with eight loudspeakers in the upper part of the figure. The lower graph visualizes the stimulus presentation concept in the time domain. Each stimulus is presented for 30 ms with 170-ms silent breaks, so the ISI is set to 200 ms.

color scheme, as explained in the color bars next to the panels. The p values are the probabilities of the null hypotheses that the distributions are significantly different (in life sciences, usually $p < 0.05$ is considered to be a significant value). The results in Figure 4.4 and 4.5 clearly show that the postulated $P300$ latency area in the range $300 - 600$ ms is the best to discriminate attended *targets* from ignored *non-targets*. This finding confirms our hypothesis that the $P300$ latencies are also related to spatial cognition in the human brain. Next in this chapter, the binary classification problem is discussed. I evaluate our hypothesis that the "handpicked" $P300$ latency ERP periods are significant features to improve the binary *target* vs. *non-target* classification accuracy. In order to find the most discriminable features from ERP responses, I used the results from the ANOVA method described above applied to the all ERP latencies. I "hand-picked" only those samples within each subject's ERPs for which the p values were smaller than 0.05 (as depicted in blue in Figure 4.4 and 4.5) in the range $300 - 600$ ms.

4.1.3.3 The Offline ERP Classification in the aBCI Paradigm

I performed the classification steps for each subject separately in aBCI offline mode, which means that all the procedures were conducted after the collection of data from each experiment, without any online feedback to the subjects. The classification procedure in our case is a so-called binary task paradigm (*target* vs. *non-target*). In the classifier training and testing step, I selected 64 *targets* and a random subset of 64 *non-targets* (from the 448 available) to have a balanced number of the members in each class set. The resulting theoretical chance level was thus 50 %. Based on our previous classification trails reported in Cai et al. [38], [44], I proposed to use a Bayesian classifier, which yielded similar or even better results on our experimental data than linear discrimination analysis methods. The GNBC is particularly suited to highly dimensional features. The GNBC method produced results comparable to more sophisticated classification methods [45] for particular cases as reported in this chapter. In our approach, I utilized a NaN-Toolbox which is a part of a BioSig environment [46]. The classifier input features were the real micro-volt EEG ERP latency values "hand-picked" as discussed in the previous section. The results of the successful application of the GNBC technique are presented in the next section.

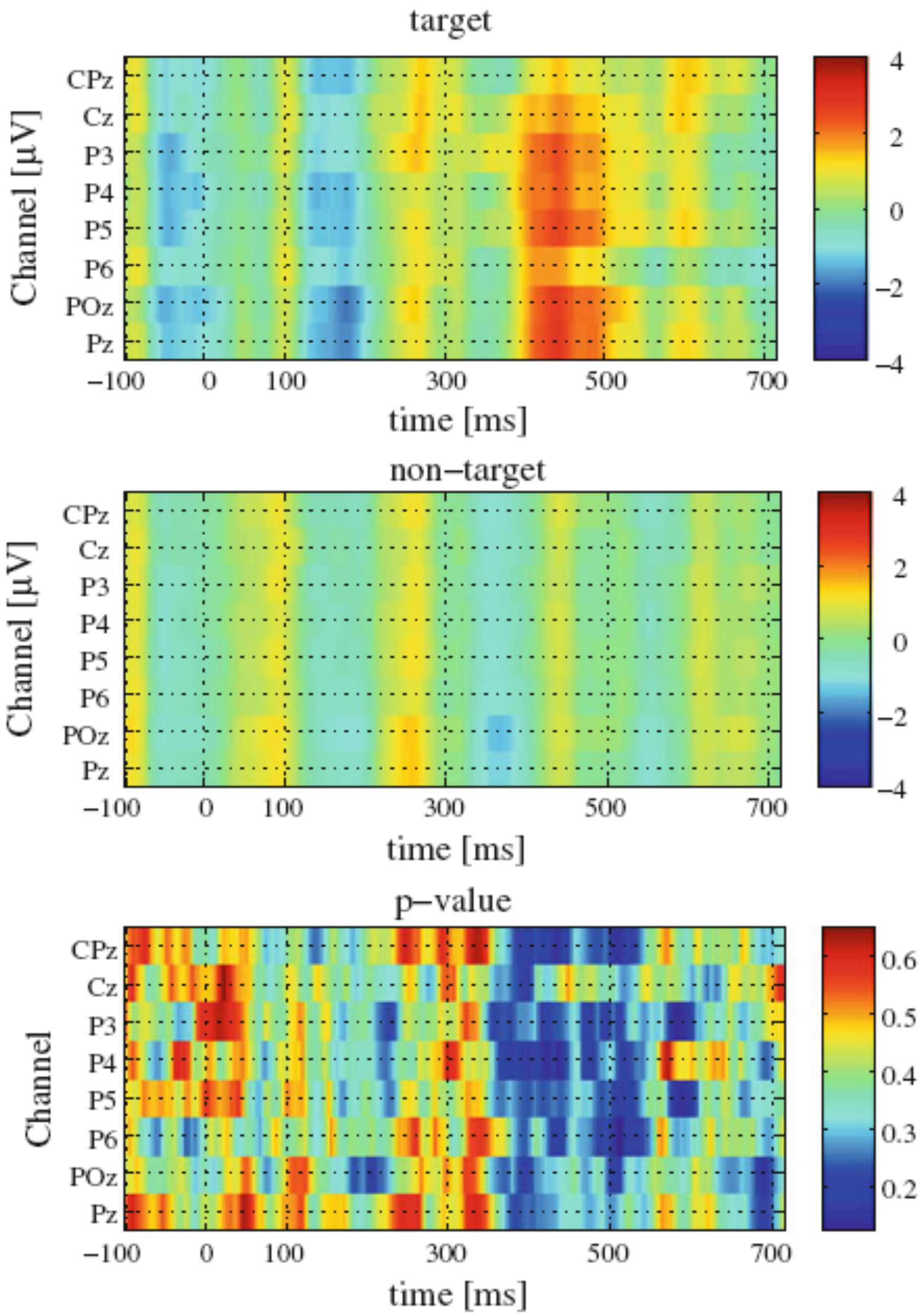

Figure 4.4: Grand mean average auditory evoked responses to spatial white noise stimuli of the ten subjects from the eight electrodes plotted separately in each row of the panels. The top panel shows the grand mean averaged response to the *targets*. The middle panel presents the grand mean averaged responses to *non-targets*. The bottom panel depicts the p values from the ANOVA for the eight electrodes separately

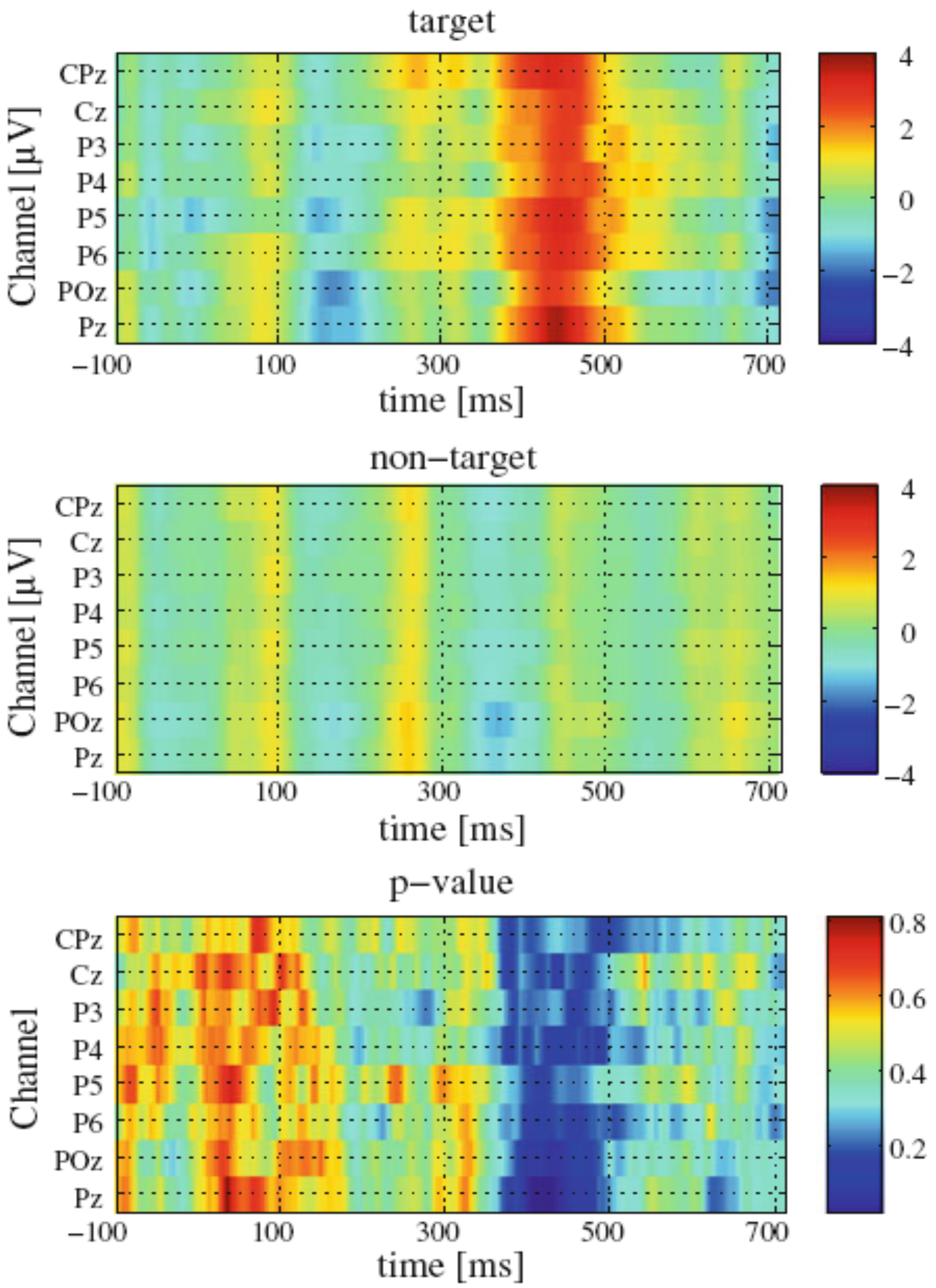

Figure 4.5: Grand mean average auditory evoked responses to spatial pink noise stimuli of the ten subjects from the eight electrodes plotted separately in each row of the panels. The top panel shows the grand mean averaged response to the *targets*. The middle panel presents the grand mean averaged responses to *non-targets*. The bottom panel depicts the p values from the ANOVA for the eight electrodes separately

4.2 Results

The proposed approach resulted in the improvement in the aBCI paradigm for setting of both the combined and averaged electrodes for each subject separately. Detailed results are presented in the following sections. First, I introduce the ITR, which is a commonly used measure to compare various paradigms in the BCI research community [12]. I also present classification accuracy results.

4.2.1 Analysis of aBCI Results with ITR and Classification Accuracies.

The amount of information carried by each selection in the BCI application is usually quantified by the ITR, which is calculated based on bits-per-selection R, defined as [43]:

$$R = \log_2 N + C \cdot \log_2 C + (1 - C) \cdot \log_2 \left(\frac{1 - C}{N - 1} \right), \tag{4.1}$$

where C is the classification accuracy and N is the number of classes ($N = 8$ in this chapter). The final ITR is obtained after a multiplication by a classification speed V, resulting in a *bit–per–minute rate* [bit/min] as:

$$ITR = V \cdot R \tag{4.2}$$

The ITR results are summarized in Tables 4.3 and 4.5 and discussed in detail in the following sections. The GNBC classification accuracies are also summarized in Tables 4.4 and 4.6.

4.2.2 The ITR and Classification Accuracy Results from the $P300$ ERP Range Latencies in the Single Channel Setting of *Target* versus *Non-target*

A summary of the ITR results is presented in Table 4.3 (the corresponding classification accuracy is summarized in Table 4.4). I compare the ITR values obtained for the features drawn from the conventional *whole-ERP* and the proposed "hand–picked" features resulting from the ANOVA. The proposed method allows us to boost the classification results up to +10.43 bit/min (+44% classification accuracy). Only a single case shows a decrease using the leave–one–out cross-validation for the GNBC technique.

4.2.3 The ITR Results from the *P*300 ERP Range Latencies from the Averaged Eight Trials in the Setting of *Target* versus *Non–target*

The single trial classification results discussed in the previous sections have resulted mostly in lower than 14 bit/min (below 70% accuracy) in the aBCI offline mode. In order to improve the results, for each subject and each stimuli, I averaged the eight *target* trails (convert 64 $targets$ to 8 $targets$) and 8 $non - target$ trials (convert 448 $non-targets$ to 56 $non-targets$). In the classifier training and testing steps, I selected 8 $targets$ and a random subset of 8 $non-targets$ (from the 56 available), the same as the single trial classification training and testing set. For each of the subjects in the case of the pink noise stimulus, the classification resulted in scores higher than 8 bit/min (80% accuracy). Especially for the subjects numbers 1 and 2, the resulting ITR reached 14.06 bit/min (100% accuracy). The summary of the results is presented in Tables 4.5 and 4.6 for ITR and classification accuracies respectively. The comparison presented of the ITR and accuracy values obtained for the features drawn from the conventional *whole-ERP* and the proposed "hand–picked" features, resulting from the ANOVA, supports the improvement of the proposed method. The method proposed allows us to increase the classification results (only a single case of a decrease was reported) using the leave–one–out cross-validation for the GNBC technique. An online aBCI application is planned as a next stage by the authors.

4.3 Discussion and Conclusions

In this chapter I have presented an approach leading to the improvement of classification accuracies and ITRs in a novel offline aBCI paradigm. This has been achieved by introducing ERP feature extraction in $P300$ range latencies to replace the classical whole evoked response range approaches.

The proposed improvement method allows the extraction of the most separable ERP features, enabling an increase in the classification accuracy and an improvement ITR of a maximum of +35.30 bit/min (22% accuracy) in the case of features drawn for single electrode ERP distributions. In the case of the features obtained from the eight trials averaged ERP responses, the majority of subjects also improved their results with a maximum increase of 10.43 bit/min (44% in accuracy).

These are the very encouraging results, providing the possibility further to improve the auditory paradigm based BCI.

The main achievement reported in the chapter allows us to improve the spatial aBCI paradigm in the offline mode, which is a step forward in non-vision based interfacing strategies. I have also shown that, in comparison with contemporary applications of spatial auditory BCI paradigms which fail to utilize rear–to–the–head loudspeakers, it is possible to utilize all spatial horizontal sound directions thanks to the proposed classification improvement approach based on the “hand–picked” ERP latencies.

Table 4.1: The spatial sound psychophysical experiment results. The response time delays and instructed directions accuracies are presented in the form of mean values with standard deviations (STD) respectively.

Pink noise stimulus								
	front-left	front-right	rear-left	rear-right	front	rear	left	right
Delay time [ms]	455	474	447	490	450	464	466	462
STD of time delay [ms]	63	76	42	75	59	76	73	65
Accuracy [%]	100	100	72	72	94	78	94	83
STD of accuracy [%]	0	0	33	25	14	27	14	28
White noise stimulus								
	front-left	front-right	rear-left	rear-right	front	rear	left	right
Delay time [ms]	439	458	448	478	445	489	450	477
STD of time delay [ms]	63	48	65	52	68	59	56	55
Accuracy [%]	100	94	83	83	94	94	94	89
STD of accuracy [%]	0	0	33	25	14	27	14	28

Table 4.2: The confusion matrix results from the psychophysical experiment averaged for all the subjects for pink– and white–noise stimulus respectively.

Confusion matrix in psychophysical tests using pink–noise stimulus								
accuracy [%]	front-left	front-right	rear-left	rear-right	front	rear	left	right
front-left	100	0	0	0	0	0	0	0
front-right	0	100	0	0	0	0	0	0
rear-left	6	6	72	5	0	0	5	0
rear-right	5	0	6	72	0	6	0	6
front	0	0	5	0	78	11	0	6
rear	0	0	5	0	6	78	0	6
left	0	0	6	0	0	0	94	0
right	0	0	0	6	0	11	0	83
Confusion matrix in psychophysical tests using white–noise stimulus								
confusion matrix accuracy [%]	front-left	front-right	rear-left	rear-right	front	rear	left	right
front-left	100	0	0	0	0	0	0	0
front-right	0	94	0	6	0	0	0	0
rear-left	0	0	83	5	0	0	0	6
rear-right	0	6	0	83	5	0	0	6
front	0	0	0	6	94	0	0	0
rear	0	0	6	0	0	94	0	0
left	0	0	0	6	0	0	94	0
right	0	0	0	6	0	11	0	83

Table 4.3: The offline aBCI interfacing results based on features drawn from non–averaged trials in the form of ITR scores obtained as in equations (4.1) and (4.2). I compare the traditional *all ERP* and the proposed "hand–picked" only latencies.

Subject	Noise stimulus type	Conventional "all ERP" [bit/min]	Proposed "hand–picked" [bit/min]	Improvement [bit/min]
#1	pink	49.39	54.13	+4.74
	white	37.90	57.42	+19.52
#2	pink	42.03	49.39	+7.36
	white	27.90	44.90	+17
#3	pink	35.26	39.25	+3.99
	white	32.72	42.03	+9.31
#4	pink	27.90	40.63	+12.73
	white	19.30	29.07	+9.77
#5	pink	47.86	66.19	+18.33
	white	47.86	57.42	+9.56
#6	pink	49.39	46.37	−3.02
	white	36.57	71.87	+35.30
#7	pink	46.37	54.13	+7.76
	white	47.86	57.42	+9.56
#8	pink	49.39	68.05	+18.66
	white	42.03	43.45	+1.42
#9	pink	50.94	50.94	0
	white	54.13	75.85	+21.72
#10	pink	39.25	37.90	−1.35
	white	44.90	49.39	+4.49

TABLE 4.4: The classification results for ERP latencies in $P300$ responses for *target* vs. *non–target* paradigm. The classification results of two feature sets (*allERP* responses and $P300$ responses) are compared. The classification improvement comparing the conventional all ERP latency with the proposed $P300$ response) is summarized in the right column.

subject	noise stimulus type	Conventional "*allERP*" [%]	Proposed $P300$ [%]	Improvement [%]
#1	pink	71	74	+3
	white	63	76	+13
#2	pink	66	71	+5
	white	55	68	+13
#3	pink	61	64	+3
	white	59	66	+7
#4	pink	55	65	+10
	white	47	56	+9
#5	pink	70	81	+11
	white	70	76	+6
#6	pink	71	69	−2
	white	62	84	+22
#7	pink	69	74	+5
	white	70	76	+6
#8	pink	71	82	+11
	white	66	67	+1
#9	pink	72	72	0
	white	74	86	+12
#10	pink	64	63	−1
	white	68	71	+3

Table 4.5: The offline aBCI interfacing results based on features drawn from the averaged eight trials in the form of ITR scores obtained as in equations (4.1) and (4.2). I compare the traditional *whole–ERP* and the proposed "hand–picked" only latencies.

Subject	Noise stimulus type	Conventional "all ERP" [bit/min]	Proposed "hand–picked" [bit/min]	Improvement [bit/min]
#1	pink	8.27	14.06	+5.79
	white	3.63	6.97	+3.34
#2	pink	6.97	14.06	+7.09
	white	4.74	8.27	+3.53
#3	pink	5.80	10.00	+4.20
	white	5.80	5.80	0
#4	pink	4.74	10.00	+5.26
	white	2.80	4.74	+1.94
#5	pink	5.80	8.27	+2.47
	white	10.00	11.74	+1.74
#6	pink	6.97	10.00	+3.03
	white	5.80	10.00	+4.20
#7	pink	3.63	14.06	+10.43
	white	6.97	8.27	+1.30
#8	pink	6.97	10.00	+3.03
	white	4.74	11.74	+7.00
#9	pink	5.8	11.74	+5.94
	white	6.97	11.74	+4.77
#10	pink	6.97	10.00	+3.03
	white	6.97	11.74	+4.77

Table 4.6: The classification results for ERP latencies in $P300$ responses for the mean of 8 *targets* vs. average of 8 *non–targets* paradigm. The classification results for two feature sets ($allERP$ responses and $P300$ responses) are compared. The classification improvement comparing the conventional all ERP latency with the proposed $P300$ response) is summarized in the right column.

subject	noise stimulus type	Conventional "$allERP$" [%]	Proposed $P300$ [%]	Improvement [%]
#1	pink	81	100	+19
	white	56	75	+19
#2	pink	75	100	+25
	white	63	81	+18
#3	pink	69	88	+19
	white	69	69	0
#4	pink	63	88	+25
	white	50	63	+13
#5	pink	69	81	+12
	white	88	94	+6
#6	pink	75	88	+13
	white	69	88	+19
#7	pink	56	100	+44
	white	75	81	+6
#8	pink	75	88	+13
	white	63	94	+31
#9	pink	69	94	+25
	white	75	94	+19
#10	pink	75	88	+13
	white	75	94	+19

Chapter 5

Utilization of $N200$ and $P300$ for Spatial aBCI Enhancement

Chapter 5 presents our results obtained with a new auditory spatial localization based BCI paradigm in which the ERP shape differences at early latencies are employed to enhance the $P300$ responses in an oddball experimental setting. The concept relies on the recent results in auditory neuroscience showing a possibility to differentiate early anterior contralateral responses to attended spatial sources. Chapter 4 BCI paradigms benefit mostly from the $P300$ ERP latencies. I show the further enhancement of the classification results in spatial auditory paradigms by incorporating the $N200$ latencies, which differentiate the brain responses to lateral, in relation to the subject head, sound locations in the auditory space. The results reveal that those early spatial auditory ERPs boost online classification results of the BCI application. The online BCI experiments with the multi–command BCI prototype support our research hypothesis could improve classification accuracies and information–transfer–rates.

5.1 Introduction

In chapter 4, I am based on $P300$ responses to distinguish *targets* and *non-targets* from series of event related potential (ERP) responses [47]. Recently a new result [48] was published elucidating the "$N200$-anterior-contralateral" (N2ac) component at the early latency (around $200ms$) of an auditory ERP. The N2ac was obtained in an experiment using two 750 ms long sound stimuli which were presented simultaneously from a different loudspeaker each. Subjects were requested to attended to the instructed *target* sound that could occur from any loudspeaker.

In this chapter, I designed a new saBCI experimental paradigm based on the auditory spatial localization principle as the informative cues with support of the N2ac component elicited in the new setup as depicted in Figure 5.1. Our hypothesis is that the new ERP component shall improve the classification results and the final information transfer rate (ITR) leading to a better BCI usage comfort.

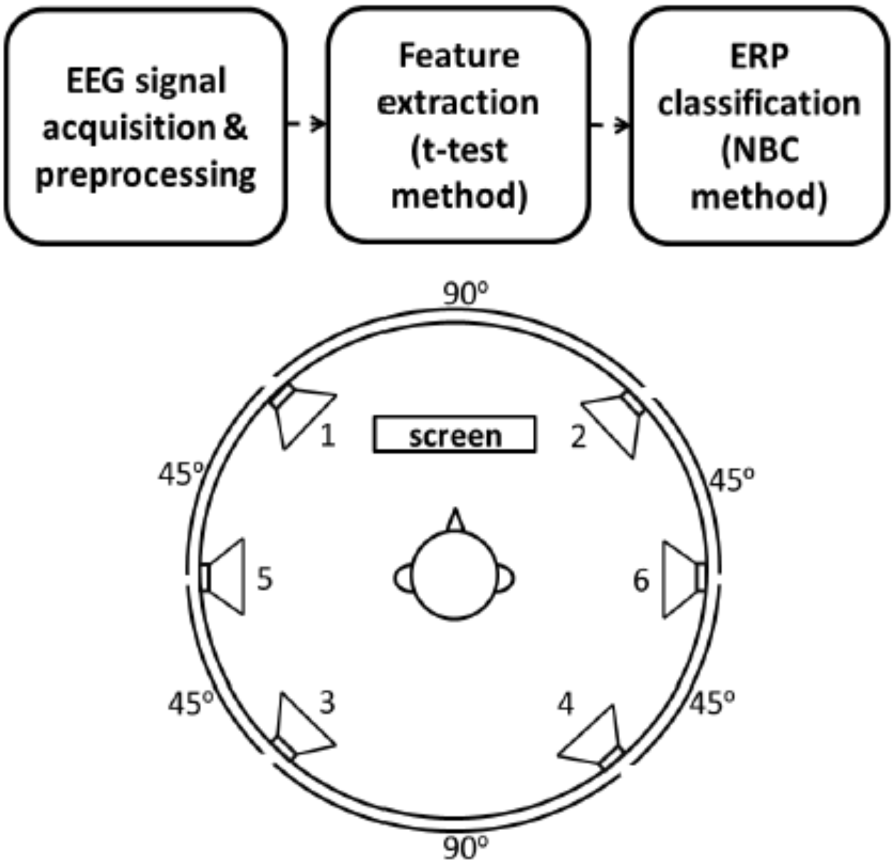

FIGURE 5.1: The novel N2apc paradigm based on spatial sound stimuli.

Within the novel saBCI paradigm framework, the subjects are asked, as in usual oddball paradigm, to attend and count the *target* stimuli from the instructed or intended direction, while ignoring the other. The EEG signals are recorded with g.MOBIlab+ EEG amplifier by g.tec. As I introduced in chapter 4, the dry g.SAHARA electrodes by the same producer which further improve the interfacing comfort, since there is no need to apply a conductive gel. In order to decrease the unnecessary and signal quality degrading muscular movement related electromyography (EMG) noise on ERP responses, the subjects are asked to minimize their eye, facial and body in general movements during the experiments. In our previous study [38] (the detail are describe in chapter 3), I proposed a channel and ERPlatency selection in order to improve the classification results. Moreover our previous study [47] (see chapter 4) ,the $P300$ response (so called "aha response" at the latency around 300 ms elicited to the expected/instructed *target* stimulus [14]) was the major feature used for the classification of the attended *targets* of the oddball paradigm. In this chapter, I introduce the early latencies around 200 ms ($N200$ response) which precede the $P300$. They shall improve the final classification rates of the saBCI application.

The objective of this chapter is to test and confirm our working hypothesis that the auditory evoked response based on N2ac paradigm should improve the saBCI application classification rates based on the new lateral to the subject head stimuli responses analysis.

From now on the chapter is organized as follows. In the next section the experimental setup and the novel paradigm is described together with EEG signals pre-processing steps. Next the analysis and optimization procedures of the ERPs at $N200$ and $P300$ response latencies for all experimental subjects are described. Finally classification and ITR results discussion conclude the chapter together with future research directions.

5.2 Method

The EEG experiments to validate the proposed spatial auditory BCI paradigm utilizing the $N200$ and $P300$ latency responses have been conducted in Multimedia Laboratory in TARA Life Science Center at the University of Tsukuba, Tsukuba, Japan. All the experimental procedure details and this approach research *targets* have been explained to the seven human subjects who agreed voluntarily to attend. The experimental procedures are designed in accordance with ethical committee guidelines of this chapter author affiliated institutions. The EEG signals are recorded by the g.USBamp EEG amplifier with the six dry g.SAHARA electrodes. The sampling frequency is set to 256 Hz with a notch filter to reject the 50Hz power line noise.

The auditory stimuli has been presented through six loudspeakers distributed with an equal radius of 1 meter around the subject's head as depicted in Figure 5.1. Three speakers with equal distances are positioned at each lateral side to the head. Two short white and pinknoise stimulus bursts are used as described in the following section. All the experiments are conducted in a silent and low reverberation room in order to limit an interference of "an environmental noise".

5.2.1 The Offine saBCI Experimental Protocol

The experimental hypothesis is that I shall be able to distinguish from the ERP shape which direction (left or right) the subject attends based on the novel N2ac response analysis method. To test the hypothesis I conduct a series of EEG recording experiments in the offine BCI mode [1] (no instant feedback or classification results given to the subject). The experiments are performed with the seven healthy subjects (six males and one female; age range $21 - -42$ with the mean of 26.4 years old). The experimental procedure has been explained in detail to each subject and her/his consent has been obtained. The subject is seated in the center of the experimental studio and the dry EEG electrodes are attached on the scalp. The subject's chair position is surrounded by the six loudspeakers. The elevation of the loudspeakers is fixed to the subject's ear level. A computer display with experimental instruction is set in front of the subject. The six loudspeakers are distributed on a circle with the three loudspeakers (1; 3; 5) positioned on the left side with 45 degrees angular distance. The remaining three loudspeakers (2; 4; 6) are located on the right side with the same angular distances (see Figure 5.1).

The sound stimulus is presented in random order one at a time from a single loudspeaker (a single trial consists of a delivery of a single *target* and five *non-targets*). As mention in chapter 4, I use two broadband noise stimulus types in order to utilize two spatial localization mechanism of the human auditory pathway (the inter-aural time and level differences - ITD/ILD) [27]. The white and pinknoise stimuli of 30 ms lengths with 5 ms linear attack and sustain periods has been chosen. The SOA is set to 300 ms. The single session consists of the six single trials (6 *targets* from each direction accompanied by 30 *non-targets*). The *target* direction in each trial is presented randomly together with five *non-targets*. For each subject and each stimuli I perform 15 sessions (all together 90 *targets* and 450 *non-targets* are delivered). The *target* direction instruction is presented visually on a computer display and auditory from the same loudspeaker which subject shall latter attend to. Before each experiment the subjects are allowed for a short practice session to familiarize themselves with spatial auditory conditions.

5.2.1.1 The Analysis of ERP Responses in Offine BCI Paradigm

In many current auditory BCI applications the focus is put on a binary classification of brain evoked responses to *targets* versus *non-targets* [12], [38], [41], [42]. The majority of the contemporary BCI applications aim at the $P300$ response latency without consideration of the remaining ERP ranges [47] (see chapter 4). Only a single of recently published papers mentions the $N200$ latency range as possibly useful to support classification [41], but there is no comparison made so far with $P300$ only related results, what I attempt in this chapter. I compare and discuss the $N200$ response suitability and I show that it really improves the classification results.

Basically a concept of adding the early latency N2ac response is based on our previous [38], [47] research and the recently published by other groups [48] concept of this ERP range modulation by ipsilateral vs. contralateral stimulus spatial locations. The ipsilateral N2ac response has higher amplitude comparing to the contralateral one. This difference confirms a feasibility to utilize the early $N200$ response latency to improve the *target* vs. nontarget classifications outcomes.

In order to precisely analyze an impact of the early ERP reposes on the saBCI paradigm classification I propose to conduct two separate analyses that shall compare how much the improvement depends only on the $N200$ response feature addition, and how much on the new feature composition based on the comparison of the ipsilateral and contralateral responses as in N2ac design.

5.2.2 EEG Preprocessing

In chapter 3 and 4, I described the EEG preprocessing part. The same processing as chapter 3 and 4. The EEG signals captured by the g.MOBIlab+ system with g.SAHARA dry electrodes are first filtered digitally with the two 5thorder Butterworth high– and low–pass filters with cutoff frequencies at 0.5 Hz and 25 Hz, respectively. Next the EEG signals are segmented creating the ERP related epochs. Each epoch starts 100 ms before each stimuli onset and ends 700 ms after it. I use the 100 ms prestimuli onset interval for a baseline correction procedures. In the next step the eye movement artifacts rejection is carried out. Auditory spatial stimuli are known to cause in subjects the uncontrolled eye movements [40] which in the current approach are removed with a threshold value set at 80μV (signal amplitude level above the usual EEG activity). The rejected epochs are not further processed, since in the current approach an emphasis is focused on the spatial paradigm validation.

5.2.3 The Optimization of the EEG Electrode Locations and ERP Features Extraction

In the previously reported research on $N2ac$ phenomenon [48] the anterior cluster of electrodes sites $F3$, $F7$, $C3$, $T7$, $F4$, $F8$, $C4$, and $T8$ was used, as in 10/20–*international system* [39]. In our experimental setup, I select the $F5$, $F6$, $C3$, $C4$, $P5$, and $P6$ electrodes (see figure 5.2) in order to have additional responses from parietal cortices known to generate ERPs related to spatial and $P300$ responses [14]. Additionally I show that the $P5$ and $P6$ sites are also useful to differentiate the responses to lateral stimuli similarly as for left–right only comparison revealed by $N2ac$. I call the new finding the $N2apc$ ($N200$–anterior–posterior–contralateral) as extension of the former one.

An example in Figure 5.3 shows the averaged and artifact–removed classical $N2ac$ responses to *ipsilateral* and *contralateral* sound stimuli as confirmed by our experiments. The presented $N200$ area responses are elucidated for *ipsilateral* and *contralateral targets*.

In order to validate statistically the differences between *target* and *non–target* responses I conduct the *t–test* analysis of the two class ERP means [49] in *ipsilateral* vs. *contralateral* experimental setting. The *t–test* method is applied to compare the differences of response distributions in single trials for each sample point of the collected ERPs. As the result I can extract discriminative information (in $N200$ and $P300$ latencies) leading to later classification optimization. The results of the

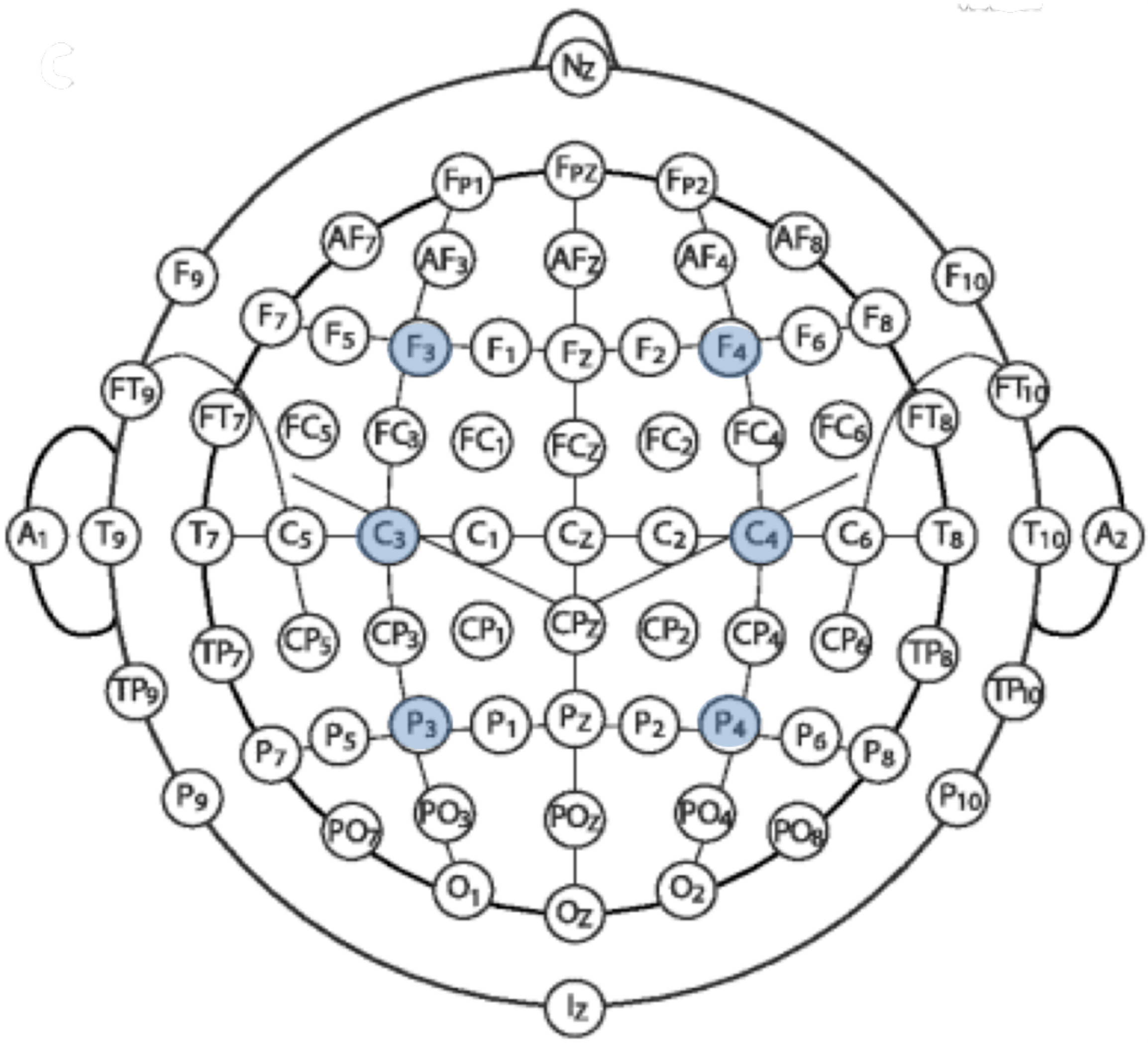

FIGURE 5.2: Six electrodes for testing N2ac on the scalp location (see blue shadow)

above analysis are depicted in Figure 5.4. A color bar located on a time scale in the above figure visualizes the *t–test's* p value results, which is a probability of the null hypothesis rejections that the means from the both compared distributions are significantly different (usually $p < 0.05$ in life sciences is considered as the significant value). The color bar in the Figure 5.4 clearly shows that the postulated $N2apc$ differential response for lateral responses is located in the range from 100 ms to 300 ms, similarly to the previously published $N2ac$ one. This finding confirms our hypothesis, that the early $N200$–range latencies are related to spatial localization processes in the human brain and that the parietal electrodes

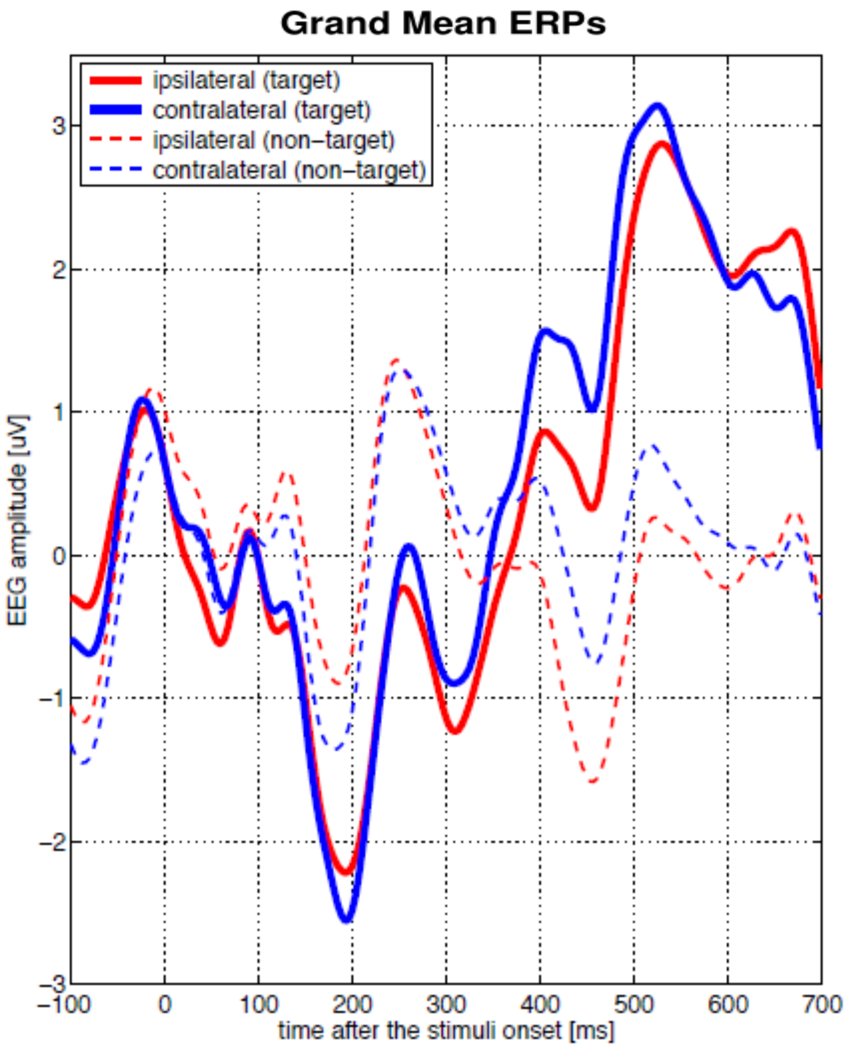

Figure 5.3: The grand mean averaged ERP responses of the seven subjects. The solid lines depict *targets* and the dashed ones *non-targets*. The red color indicates ipsilateral and blue one the contralateral responses. The differences between *targets* and *non-targets* are obvious after 300 ms (the so called "aha" or $P300$ response), while the lateral directions can be identified in $N200$ latency area.

contribute also to the result.

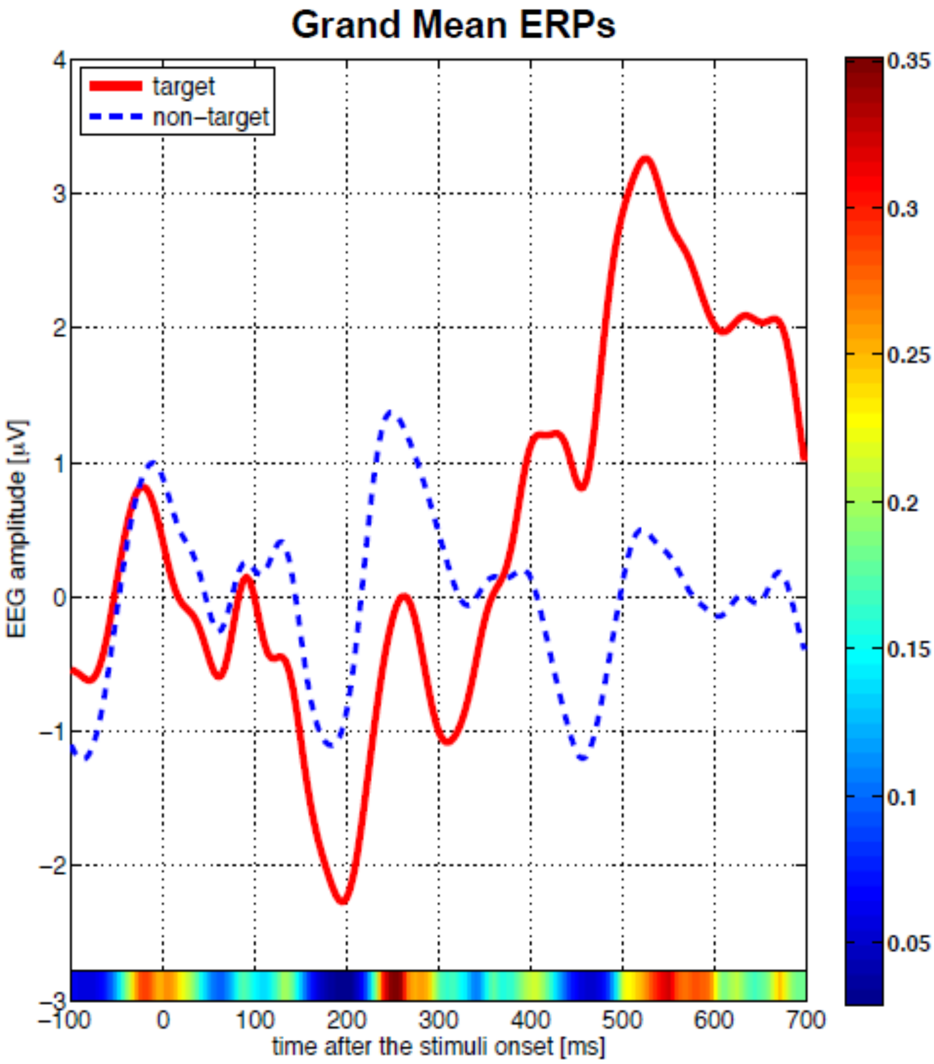

Figure 5.4: The grand mean averaged ERP for the all seven subjects and all electrodes calculated together, while plotted separately for *target* (solid red line) and non–target (dashed blue line) responses. The significant differences between the both responses can be found, as visualized by the color bar with p–values of t–test results (statistical significance for $p < 0.05$) in the bottom part in the above panel, can be found around 200 ms ($N200$ response latency) and after 300 ms ($P300$ response latency).

In this chapter two types of binary classification problems are discussed. First I evaluate our first hypothesis that adding the early latency ERP periods as features improves the binary *target* vs. *non-target* classification. Next I also show that the novel *N2apc* response further enhances the results using the *ipsilateral* vs. *contralateral* response comparison.

In order to find the most discriminable features from ERP responses I use the results of the above described *t–tests* evaluating statistical significance of them. I "hand pick" only those samples within each subject's ERPs for which the p–values are smaller than 0.05 as depicted by blue shades of the color bar at the bottom of the Figure 5.4. The significantly different ERP samples of $N2apc$ based experiments (I relax here the condition to $p < 0.10$ only) are depicted in Figures 5.5 and 5.6 for pink– and white–noise stimuli respectively. In the next section I show that the relaxed condition of *t–test's* $p < 0.10$ improves already satisfactory the saBCI classification results by incorporating the $N200$ latency responses.

5.2.4 The Offline saBCI Classification

I perform the classification steps for each subject separately in saBCI offline mode, which means that all procedures are conducted after each experiment of data collection, without any online feedback to subjects. The classification procedure is performed in a so called binary task paradigm (I classify *target* vs. *non–target*, or *contralateral* vs. *ipsilateral* response pairs each time only).

In each classifier training and testing step I select 90 *targets* and a random subset of 90 *non-targets* (from the 450 available) to have the balanced number of the members in each class set. The resulting chance level is 50%. For the case of the *contralateral* vs. *ipsilateral* responses classification I select 30 *contralateral* and 30 *ipsilateral* events.

Based on our previous classification trials reported in [38] I decide to use a Baysian classifier, which outperforms the linear discrimination analysis methods. *The naive-Bayses classifier* (NBC) is particularly suited for the highly dimensional features. Despite its simplicity, the NBC approach often outperforms more sophisticated classification methods [34]. The NBC application assigns an unknown sample (ERP features in our case) $\mathbf{x} = [x_1, x_2, \ldots, x_l]^T$ based on probability maximization to the class

$$\omega_m = \arg\max_{\omega_i} \prod_{j=1}^{l} p(x_j|\omega_i), \quad i = 1, 2, \ldots, M, \tag{5.1}$$

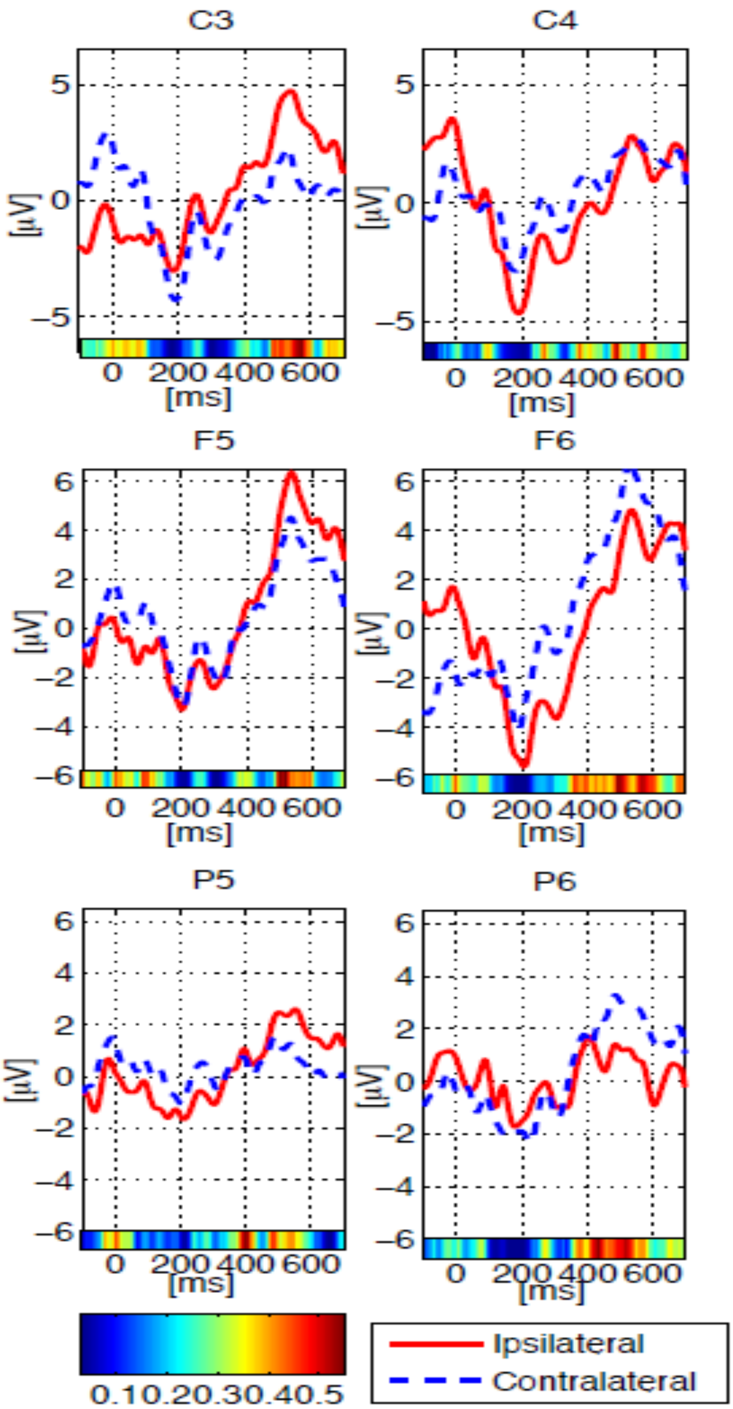

FIGURE 5.5: ERP to pink noise stimuli grand mean averages for all subjects and the six electrodes plotted separately in each panel. The solid red lines represent the ipsilateral to *target* responses and the dashed blue lines to the contralateral ones. The color bars at the bottom of each panel show the *t-test* resulting $p-$values.

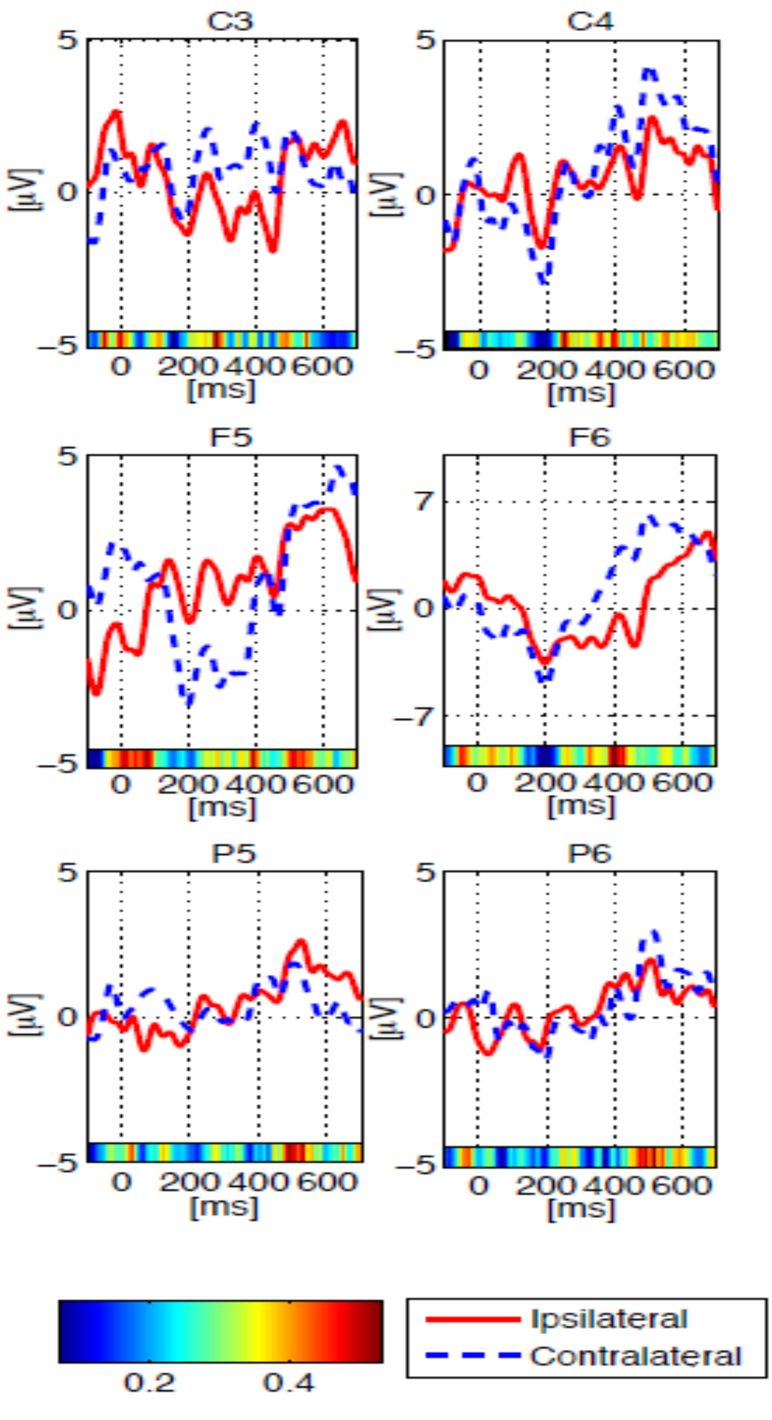

FIGURE 5.6: ERP to white noise stimuli grand mean averages for all subjects and the six electrodes plotted separately in each panel. The solid red lines represent the ipsilateral to *target* responses and the dashed blue lines to the contralateral ones. The color bars at the bottom of each panel show the *t-test* resulting $p-$ values.

with an assumption that the individual features x_j, $j = 1, 2, \ldots, l$, shall be statistically independent. It turns out that the NBC can be very robust also to violations of the independence assumption [34].

Consider the vector $\mathbf{x}$ with features according to the values of the ERP "hand picked" samples. The respective conditional probabilities shall be $P(x_i|\omega_1) = p_i$ and $P(x_i|\omega_2) = q_i$, in our binary classification case comparing *targets* vs. *non-targets* or *ipsilateral* vs. *contralateral* responses. In Bayesian rule, given the value of $\mathbf{x}$ the class membership is decided according to the probabilities likelihood ratio

$$\frac{P(\omega_1)P(\mathbf{x}|\omega_1)}{P(\omega_2)P(\mathbf{x}|\omega_2)} > (<)1. \tag{5.2}$$

The adoption of features independence principle allows us to limit a number of necessary training samples and I can write

$$P(\mathbf{x}|\omega_1) = \prod_{i=1}^{l} p_i^{x_i}(1-p_i)^{1-x_i} \tag{5.3}$$

$$P(\mathbf{x}|\omega_2) = \prod_{i=1}^{l} q_i^{x_i}(1-q_i)^{1-x_i} \tag{5.4}$$

Now an application of a logarithm function to the both sides of the equation (5.2) results with a linear discriminant function as

$$\begin{aligned} h(\mathbf{x}) &= \sum_{i=1}^{l}\left(x_i \ln\frac{p_i}{q_i} + (1-x_i)\ln\frac{1-p_i}{1-q_i}\right) \\ &+ \ln\frac{P(\omega_1)}{P(\omega_2)}, \end{aligned} \tag{5.5}$$

which could be brought to the linear form of

$$h(\mathbf{x}) = \mathbf{w}^T\mathbf{x} + w_0, \tag{5.6}$$

based on the following substitutions

$$\begin{aligned} \mathbf{w} &= \left[\ln\frac{p_1(1-q_1)}{q_1(1-p_1)}, \ldots, \ln\frac{p_l(1-q_l)}{q_l(1-p_l)}\right]^T \\ w_0 &= \sum_{i=1}^{l}\ln\frac{1-p_i}{1-q_i} + \ln\frac{P(\omega_1)}{P(\omega_2)}. \end{aligned}$$

The results of NBC technique successful application are presented in the next section.

5.3 Results

As the result of the presented research have obtained the results showing that for the both experimental settings of saBCI offline paradigm the classical $P300$ latency could be improved with the pure $N200$ or the more complex $N2apc$ features identified with p–values calculated using the classical *t–test* for significance. I summarize below the obtained results.

5.3.1 The Classification Results from the Combined $N200$ and $P300$ ERP Latencies in the Classical *target* vs. *non–target* Setting

The first summary of classification results is presented in Table 6.1, where classification accuracies for the features drawn from $N200$, $P300$ and the combined latencies are shown. The majority of the subjects performed already above the chance level of 50% (except subject MA for the pink noise case) for single feature latencies of $N200$ or $P300$. The proposed combination of the two "hand–picked" feature sets using the *t–test* significant ERP samples allowed us to boost the classification results up to 7% (only a single case of the accuracy decrease has been reported) using the *leave–one–out* cross validation [34] for the NBC technique.

5.3.2 The Classification Results from the new $N2apc$ ERP Feature in the Ipsilateral vs. Contralateral Settings

The results of the proposed approach to compare *ipsilateral* and *contralateral* to target evoked potentials have been summarized in the Table 6.2, based on the ERP features drawn from results of the *t-test* analysis as summarized in the Figures 5.5 and 5.6. The classification accuracy results have been 17% boosted in the best case, with the same method of the NBC *leave–one–out* cross validation.

5.3.3 Analysis of Information Transfer Rate Improvement Results

The amount of information carried by every selection in the BCI application is usually quantified by the ITR which is calculated based on bits–per–selection R, defined as in [43]. The detail description of ITR is presented in chapter 4. The ITR results are summarized in Tables 5.3 and 5.4. For the both cases of the $N200/P300$ combination and the $N2apc$ paradigm, there is a significant increase of ITR for the majority of subjects.

5.4 Conclusions

In this chapter I presented two approaches leading to improvements of classification accuracy and ITR in offline saBCI paradigm by introducing the novel ERP feature extraction in combined $N200/P300$ latencies and in the new $N2apc$ setting which compares responses of lateral, to the head, sound sources.

The first improvement analysis resulted in a comparison of classification rates for the three ERP feature sets of $N200$ and $P300$ latencies processed separately, versus the combined $N200/P300$. The latter combination resulted in a steady increase in classification accuracy for the majority of subjects up to 7% at maximum. Additionally the ITR improvement in this case was reported at maximum of 7bit/min. This is a very good result giving a possibility to further improve the auditory paradigm based BCI.

The second improvement step is based on the proposed extension of $N2ac$ concept. I added a comparison of parietal electrodes responses allowing for the new feature creation from such ERP comparisons. The new ERP component was named $N2apc$ since it combines anterior and posterior contralateral response differences. The obtained classification and ITR improvement was also very encouraging.

The two main achievements reported in the chapter allowed us to improve the novel saBCI paradigm in offline mode which is a step forward in the non–vision based interfacing strategies. The obtained results reveal that not only the cortical auditory information processing centers related to the cognitive streams could be utilized to BCI purposes. Also the differences in ERPs at early latencies before

300 ms are useful and they guarantee good classification results and ITRs. These results reveal that the very early spatial auditory ERPs are potentially interesting for faster BCI applications.

Table 5.1: The classification results for ERP latencies in $N200$ and $P300$ responses for *target* vs. *non–target* paradigm. The three feature sets ($N200$, $P300$ and $N200/P300$ latencies combined) classification results are compared. The classification improvement comparing the classical $P300$ latency only with the proposed combination of $N200/P300$) is summarized in the right column.

subject	noise stimulus type	$N200$ only [%]	$P300$ only [%]	$N200/P300$ combined [%]	$N200/P300$ vs. $P300$ [%]
#1	pink	63	63	64	1
	white	54	59	60	1
#2	pink	52	54	56	2
	white	56	69	68	−1
#3	pink	53	57	57	0
	white	57	57	58	1
#4	pink	60	69	69	0
	white	55	58	65	7
#5	pink	65	65	67	2
	white	46	40	44	4
#6	pink	54	59	59	0
	white	53	52	53	1
#7	pink	64	61	66	5
	white	53	61	63	2

Table 5.2: The classification results for the proposed method using *N2apc* response to support the saBCI compared with the conventional method.

subject	noise stimulus type	conventional method [%]	N2apc paradigm [%]	the improvement [%]
#1	pink	56	61	5
	white	51	63	12
#2	pink	52	61	9
	white	63	67	4
#3	pink	58	58	0
	white	47	54	7
#4	pink	37	54	17
	white	49	50	1
#5	pink	49	54	5
	white	50	56	6
#6	pink	32	48	16
	white	48	50	2
#7	pink	72	68	−4
	white	58	69	11

TABLE 5.3: The ITR for the three ERP interval related classification approaches using $N200$ or $P300$ only, and the combined $N200/P300$ together.

subject	noise stimulus type	$N200$ only [bit/min]	$P300$ only [bit/min]	$N200/P300$ combined [bit/min]	$N200/P300$ vs. $P300$ [bit/min]
#1	pink	25.84	25.84	26.88	1.04
	white	17.38	21.88	22.84	0.96
#2	pink	15.72	17.38	19.12	1.74
	white	19.12	32.40	31.25	−1.15
#3	pink	16.14	20.02	20.02	0.00
	white	22.84	20.02	20.94	0.92
#4	pink	22.84	32.40	32.40	0.00
	white	18.24	20.94	27.94	7.00
#5	pink	27.94	27.94	30.13	2.19
	white	11.19	7.36	9.84	2.48
#6	pink	17.38	21.88	21.88	0.00
	white	16.54	15.72	16.54	0.82
#7	pink	26.88	23.82	29.02	5.20
	white	16.54	23.82	25.84	2.02

TABLE 5.4: The ITR for the proposed method using $N2apc$ response to support the saBCI classification rates.

subject	noise stimulus type	conventional method [bit/min]	proposed N2ac [bit/min]	resulting change [bit/min]
#1	pink	19.12	23.82	4.70
	white	14.92	25.84	10.92
#2	pink	15.72	23.82	8.10
	white	25.84	30.13	4.29
#3	pink	20.94	20.94	0.00
	white	11.90	17.38	5.48
#4	pink	5.72	17.38	11.66
	white	13.37	14.13	0.76
#5	pink	13.37	17.38	4.01
	white	14.13	19.12	4.99
#6	pink	3.39	12.62	9.23
	white	12.62	14.13	1.51
#7	pink	35.98	31.25	−4.73
	white	20.94	32.4	11.46

Chapter 6

ERP Responses to Front–Back to the Head Stimuli Distinction Support

Chapter 4 shows a possibility to differentiate early anterior contralateral responses to attended spatial sources. I also found that the early brain responses elucidate which direction, front or rear loudspeaker source, subject attended. I show the further enhancement of the classification results in spatial auditory paradigm, in which I incorporate the $N200$ latencies. The results reveal that those early spatial auditory ERPs boost offline classification results of the BCI application. The offline BCI experiments with the multi-command BCI prototype support our research hypothesis with the higher classification results and the improved information–transfer–rates.

6.1 Introduction

The majority of the contemporary BCI applications aim at the $P300$ response latency without consideration of the remaining the ERP ranges (see chapter 3 and 4). In this chapter, I compare and discuss the $N200$ response suitability and I show that its utilization improves final classification results.

As a first step I propose to utilize the early ERP latency based modulation related to the so called "$N200$–anterior–contralateral" ($N2ac$) response [48] (see chapter 5). This response is characterized by different shapes in the brain ipsilateral and contralateral to *targets* ERPs. Our previous research also confirmed previously [44] that the $N2ac$ setting improved the BCI classification accuracy. Using this setting allowed us to classify the *targets* from left and right side loudspeaker *targets* respectively. Psychophysical experiments also confirmed that the subjects had no discrimination problems related to the auditory *front–back–confusion*. The other previous study, conducted by the authors, further confirmed the feasibility to utilize front and rear loudspeaker directions for the saBCI paradigm [38].

In this chapter, I report on the new finding that the different ERP shape at the $N200$ latency support discrimination between attended auditory stimulus *targets* originating from frontal and rear loudspeakers respectively. This new finding allows us to identify a auditory stimulus direction to which the subject attended (front or rear). I call this new finding "$N200$–front–rear" ($N2fr$).

I report also on a design of a new saBCI experimental paradigm based on the auditory spatial localization principle as the informative cue with support of the both $N2ac$ and $N2fr$ ERP components elicited in the new experimental setup as depicted in Figure 6.1. Our hypothesis is that the new ERP component shall improve the classification results and the final information transfer rate (ITR) leading to a better BCI usage comfort in general.

6.2 Methods

The EEG experiments to validate the proposed spatial auditory BCI paradigm utilizing the $N200$ and $P300$ latency responses were conducted in Multimedia Laboratory in Life Science Center of TARA at the University of Tsukuba, Tsukuba, Japan. All the experimental procedure details and this approach research *targets* were explained to the seven human subjects who agreed voluntarily to attend. The experimental procedures were designed in accordance with ethical committee guidelines of this chapter author affiliated institutions. The EEG signals were recorded by the g.USBamp EEG amplifier with the six dry g.SAHARA electrodes. The sampling frequency was set to 256 Hz with a notch filter to reject the 50Hz power–line noise set to remove EEG signal band on a range of $48 \sim 52$Hz.

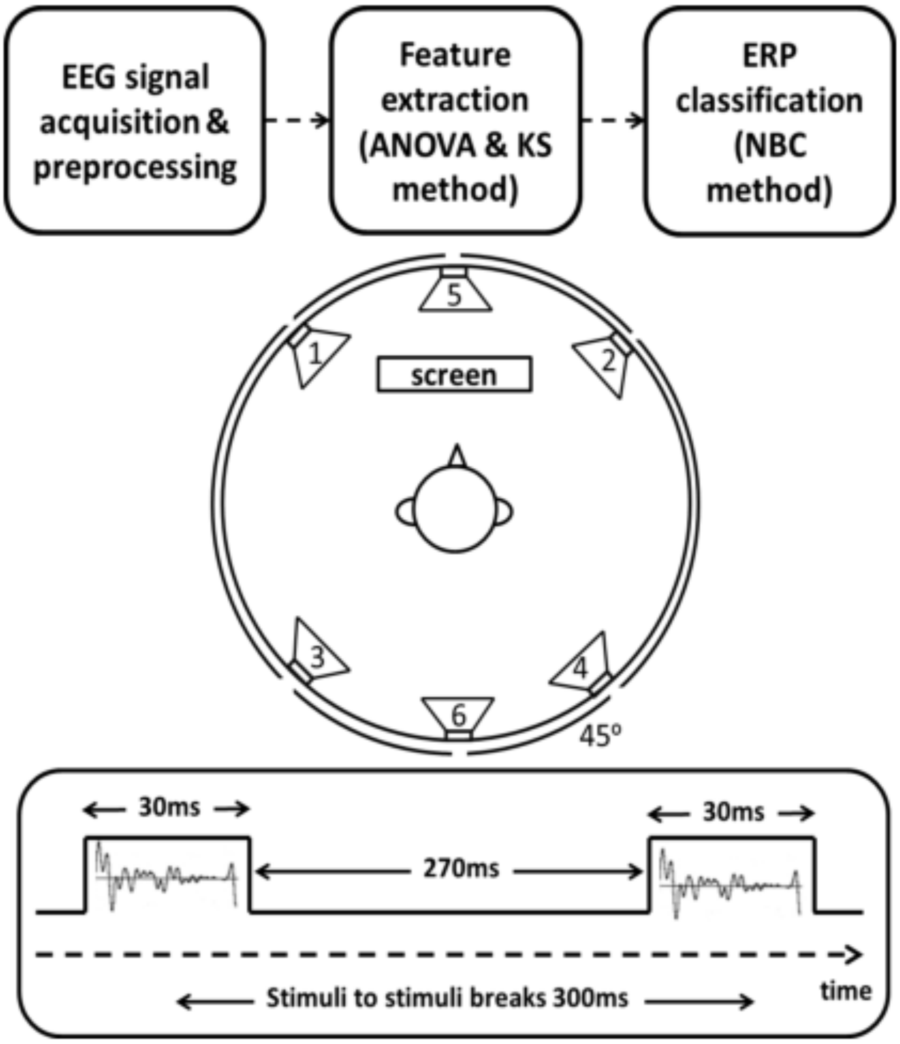

Figure 6.1: The novel front and rear to the subject head's auditory sources localization paradigm based on spatial sound stimuli is depicted in the top panel. The bottom panel presents our stimulus presentation concept illustrated in the time domain. Each stimulus has been presented in our experiments for 30 ms with 270 ms silent breaks with the respective inter–stimulus–interval (ISI) of 300 ms.

The auditory stimulus was presented through the six loudspeakers distributed with an equal radius of one meter around the subject's head as depicted in Figure 6.1. Three speakers with equal distances were positioned at frontal and rear sides to the head. Two short *white–* and *pink–*noise stimulus bursts were used as described in the following section. All the experiments were conducted in a silent and low reverberation room in order to limit environmental noise interferences.

6.2.1 The Offline saBCI Experimental Protocol

The experimental hypothesis was that I shall be able to distinguish from the ERP shape which direction left, right, front or rear to the head the subject attended based on the $N2ac$ and $N2fr$ responses.

To test the hypothesis I conducted a series of EEG recording experiments in the offline BCI mode [1] (no instant feedback or classification results given to the subject). The experiments were performed with the seven healthy subjects (six males and one female; age range $21 - 42$ with the mean of 26.4 years old). The experimental procedure was explained in detail to each subject and her/his written consent was obtained. The subject seated in the center of the experimental studio and the dry EEG electrodes were attached on the scalp. The subject's chair was positioned in the middle of the surrounding six loudspeakers. The elevation of the loudspeakers was fixed to the subject's ear level. A computer display with experimental instruction was set in front of the subject. The six loudspeakers were distributed on a circle with the three loudspeakers $(1, 2, 5)$ positioned in the front with 45° angular distance. The remaining three loudspeakers $(3, 4, 6)$ were located in the rear with the same angular distances (see Figure 6.1). The four loudspeakers $(1, 2, 3, 4)$ were used to test the $N2ac$ effect, and two other loudspeakers $(5, 6)$ were used to confirm our hypothesis of $N2fr$ response.

The sound stimulus was presented in a random order, one at a time from a single loudspeaker (a single trial consisted of a delivery of the single *target* and five *non-targets*). I decided to use two broadband noise stimulus types in order to utilize two spatial localization mechanism of the human auditory pathway (the inter-aural time and level differences – ITD/ILD) [27]. The *white–* and *pink–*noise stimuli of 30 ms lengths with 5 ms linear attack and sustain periods were chosen. The inter–stimulus–interval (ISI) was set to 300 ms. The single session consisted of the six single trials (6 *targets* from each direction accompanied by 30 non-targets).

The *target* direction in each trial was presented randomly together with five non-targets. For each subject and each stimuli I performed 15 sessions (all together 90 *targets* and 450 *non-targets* were delivered). The *target* direction instruction was presented visually on a computer display and auditory from the same loudspeaker which subject shall latter attend to. Before each experiment the subjects were allowed for a short practice session to familiarize themselves with spatial auditory conditions.

6.3 The Analysis of ERP Responses in Offline BCI Paradigm

In many current auditory BCI applications the focus is put on a binary classification of brain evoked responses to *targets* versus non-targets [38, 41–43]. The majority of the contemporary BCI applications aim at the $P300$ response latency without consideration of the remaining ERP ranges. Only a single of recently published papers mentioned the $N200$ latency range as possibly useful to support classification [41], but there is no comparison made so far with $P300$ only related results, what I attempt in this chapter.

Basically a concept of adding the early latency $N2ac$ or $N2fr$ responses is based on our previous [38] research and the recently published by other groups [48] concept of this ERP range modulation by *ipsilateral* vs. *contralateral* stimulus spatial locations, which results in the different ERP shapes. This difference confirms a feasibility to utilize the early $N200$ response latency to improve the *target* vs. *non–target* classification accuracy.

In order to precisely analyze an impact of the early ERP reposes on the saBCI paradigm classification I propose to conduct three separate analyses that shall compare results from:

- the $P300$ only based BCI classification;
- the $N2ac$ response based improvement;
- the further $N2fr$ related classification boosting.

6.3.1 EEG Preprocessing

The EEG signals captured by the g.MOBIlab+ system with g.SAHARA dry electrodes were first filtered digitally with the two 5^{th}–order Butterworth high– and low–pass filters with cutoff frequencies of 0.5 Hz and 25 Hz, respectively.

The high–pass filtering removed the very slow baseline drift related artifacts as well as the slow eye movements related EMG interferences. The low–pass filter limited the higher frequency EMG artifacts related to subject body muscle movements.

Next the EEG signals were segmented creating the ERP related *epochs*. Each *epoch* started 100 ms before each stimuli onset and ended 700 ms after it. I used the 100 ms pre–stimuli onset interval for baseline correction procedures.

In the next step the eye movement artifacts rejection was carried out. Auditory spatial stimuli has been known known to cause in subjects the uncontrolled eye movements [40] which in the current approach were removed with a threshold value set at 80μV (signal amplitude level above the usual EEG activity). The rejected epochs were not further processed, since in the current approach an emphasis was focused on the spatial paradigm validation. An example in Figure 6.2 shows the averaged and artifact–removed $P300$ responses to *target* and *non-target* sound stimuli with standard error bars respectively.

6.3.2 The Optimization of the EEG Electrode Locations and ERP Feature Extraction

In the previously reported research on $N2ac$ phenomenon [48] the anterior cluster of electrodes sites $F3$, $F7$, $C3$, $T7$, $F4$, $F8$, $C4$, and $T8$ was used, as in 10/20–*international system* [39]. In our experimental setup, I select the $F5$, $F6$, $C3$, $C4$, $P5$, and $P6$ electrodes in order to have additional responses from parietal cortices known to generate ERPs related to spatial and $P300$ responses [14]. Additionally I show that the $P5$ and $P6$ sites are also useful to differentiate the responses to lateral stimuli similarly as for left vs. right only comparison revealed by $N2ac$.

An example in Figure 6.4 shows the averaged and artifact–removed $N2ac$ responses to *ipsilateral* and *contralateral* sound stimuli as confirmed by our experiments. The presented $N200$ area responses are elucidated for *ipsilateral* and *contralateral targets*.

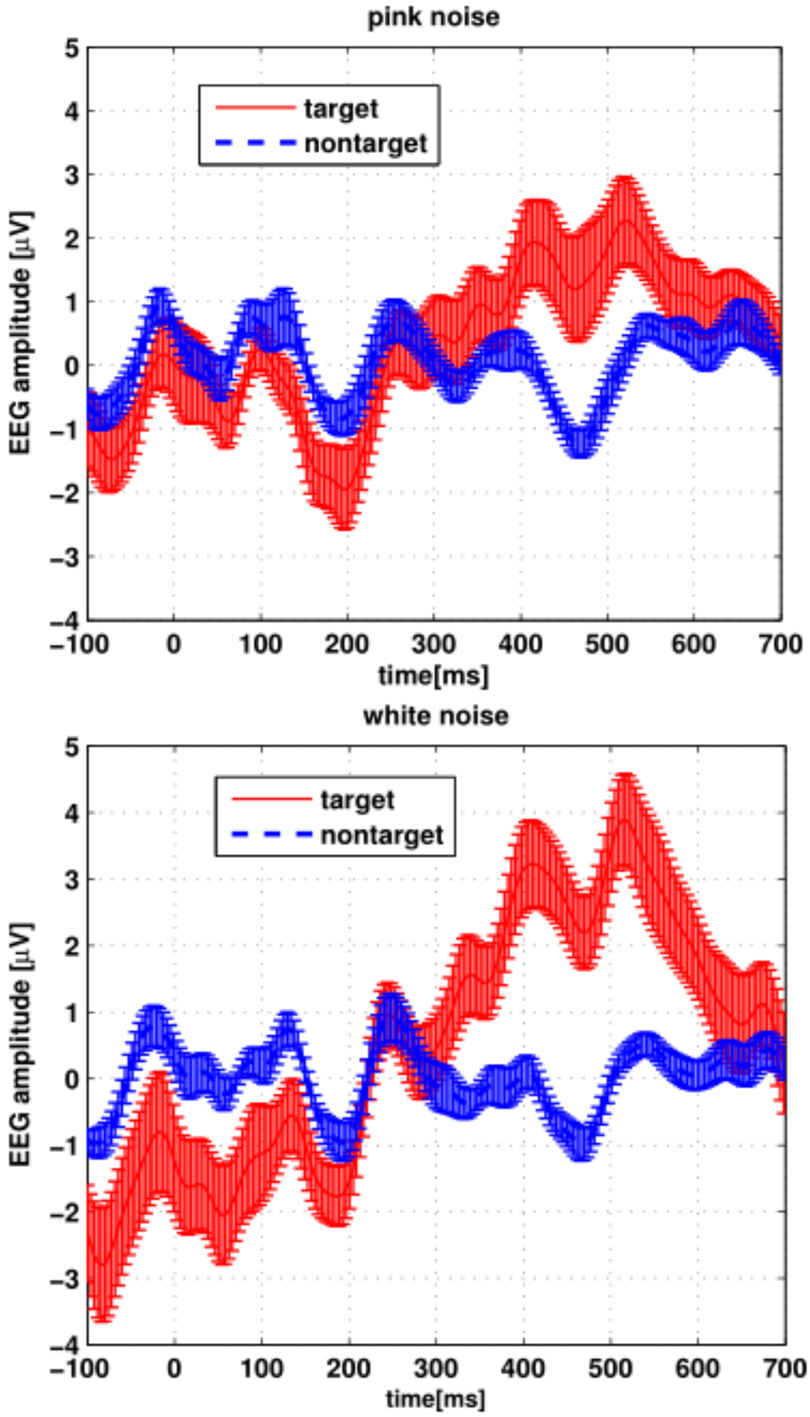

Figure 6.2: The grand mean averaged ERP responses of the seven subjects. The upper panel present the grand mean averaged ERP responses to pink–noise. The lower panel depicts respective results obtained with white–noise stimulus. The red lines represent *targets* and the blue ones non-targets. All results are presented together with standard error bars. The differences between *targets* and non-targets are obvious in the range of 300 $\sim$ 600 ms (the so called "aha"– or $P300$ response).

In order to validate statistically the differences between *target* and *non–target* responses I conduct the analysis of variance test (ANOVA) [49] and Kolmogorov–Smirnov test (KS) (to compare the two class ERP distribution similarities). I also used those methods to analyze the two class ERP distributions in *ipsilateral* vs. *contralateral*, as well as *the frontal–* vs. *rear– loudspeaker* originating *targets* respectively. The ANOVA– and KS–test methods were applied to compare the differences of response distributions in single trials for each sample point of the collected ERPs. As the result I could extract discriminative information (in $N200$ and $P300$ latencies) leading to later classification optimization. The results of *targets* and *non-targets* analyses are depicted in Figure 6.3. The both bottom panels in the above figure visualize the ANOVA– and KS–test $p-$value results, which are the probability of the null hypothesis rejections that the distributions from the both compared ERPs are significantly different (usually $p < 0.05$ in life sciences is considered as the significant value). The ANOVA–test results panel the Figure 6.3 shows that the $P300$ related significant responses are located in the range from 300 ms to 500 ms. The ERP analysis results of ipsilateral vs. contralateral responses is shown in Figure 6.4. The $p-$values obtained from ANOVA– and KS–tests and presented in Figure 6.4 depict that the $N200$ singificant responses are located in the range from 100 ms to $300ms$ similarly to the previously published $N2ac$ phenomenon. This finding confirms our hypothesis that the early $N200$–range latencies are related to spatial localization processes in the human brain and that the parietal electrodes contribute also to the result. The ERP analysis results of the $N2fr$ experimental setting are depicted in Figure 6.5. The ANOVA– and KS–tests resulting $p-$values confirm that the significant $N200$ response is located in the range form 100 ms to 200 ms. This difference is a source of the next discussed in the chapter BCI classification accuracy improvement.

In this chapter three types of binary classification related feature extraction problems have beed outlined so far:

1. First, I have evaluated ERP features separability in the $P300$ latency for the classical *target* vs. *non-target* classification purposes.

2. Second, I have shown that the $N2ac$ latency further have contribute to the ERP features separability in the *ipsilateral* vs. *contralateral* stimulus presentation setting.

3. Finally, I have postulated and validate the new concept of *frontal–* vs. *rear–loudspeaker* stimulus response differences which I named the $N2fr$.

I propose to identify the most discriminable features from ERP responses using the above described ANOVA– and KS–tests evaluating the statistical significance of the latency differences. Next “a hand picking” procedures could be applied of only those ERP latencies within each subject ERPs p–values are smaller than a reasonable threshold of around 0.20 as depicted by blue shades in the third from the top and bottom panels in Figures 5.4, 6.4 and 6.5.

In a next section I show that the proposed method of utilizing ANOVA– and KS–tests for $p < 0.20$ improves the saBCI classification results with utilizing the $N200$ latency responses.

6.3.3 The Offline saBCI Classification

In order to test the proposed feature extraction method the offline saBCI mode classification have been performed for each subject separately, meaning that all procedures have been conducted after each experiment of data collection, without any online feedback to subjects. The classification procedure has been performed in a so called binary task paradigm of *targets* vs. *non-targets*, *contralateral* vs. *ipsilateral*, or *frontal–* vs. *rear–loudspeaker targets* respectively.

In each classifier training and testing step 90 *targets* and a random subset of 90 *non-targets* have been selected (from the 450 available) to have the balanced numbers in the each class set. The resulting chance level was 50%.

For the case of the *contralateral* vs. *ipsilateral* responses classification I selected 30 *contralateral* and 30 *ipsilateral* events.

To classify the *targets* from frontal– vs. rear–loudspeakers I selected 15 *targets* from each direction respectively.

Based on our previous classification trials reported in [44] I decide to continue to use a Baysian classifier, which for our case datasets outperforms the linear discrimination analysis methods. Despite its simplicity, the NBC (naive Bayesian classifier) approach often outperforms more sophisticated classification methods [34]. The NBC application assigns an unknown sample (ERP features in our

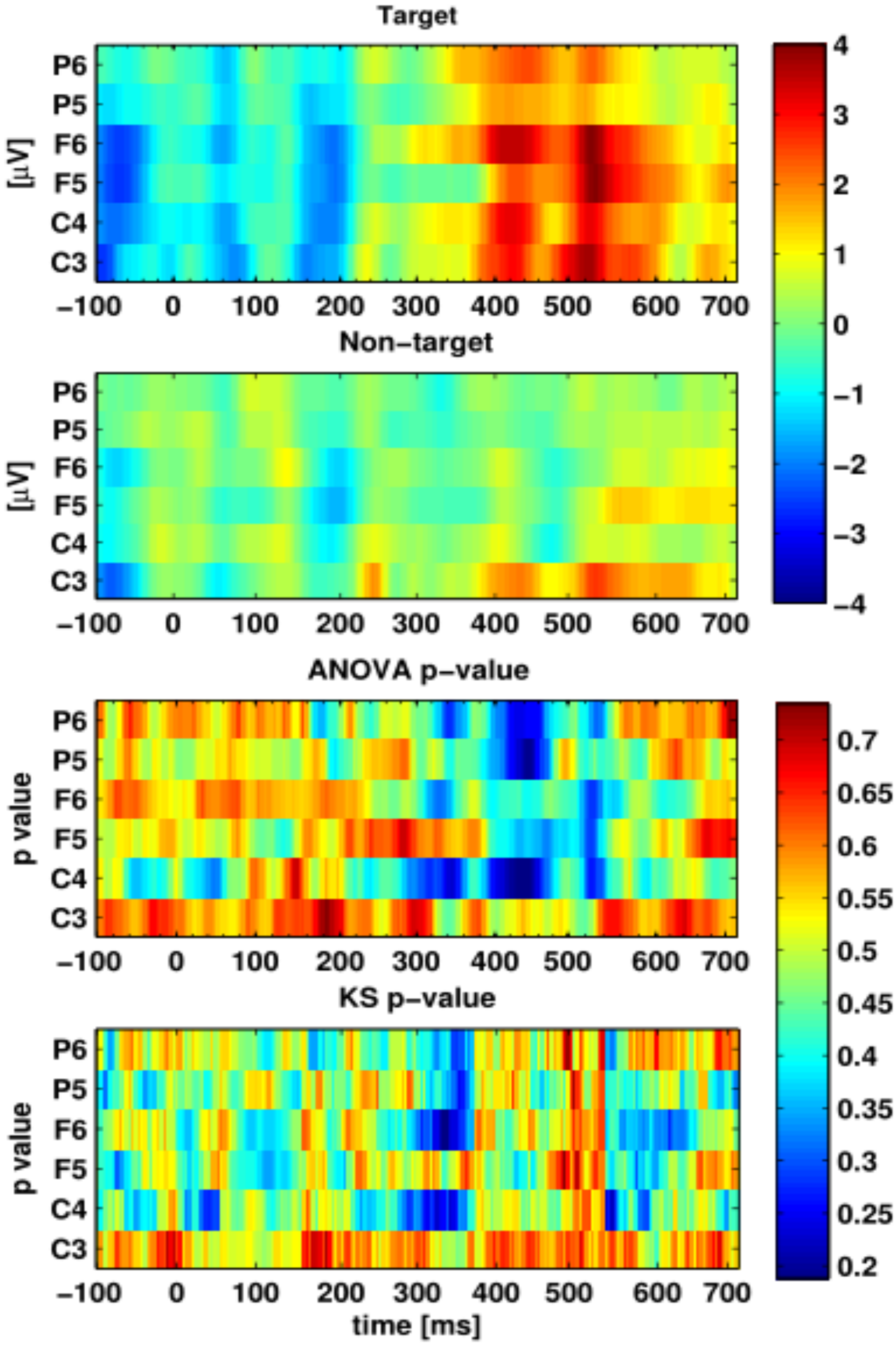

FIGURE 6.3: Grand mean averaged ERP responses of the all seven subjects and the six electrodes are plotted separately in the two top panels. The first top panel shows *target* and a second from the top the *non–target* averaged responses, respectively. The significant differences between the both responses can be found, as visualized by color coding of the p-values obtained from *ANOVA–test* (statistical significance at $p < 0.05$) in the third from the top panel. The bottom panels presents *Kolmogorov–Smirnov test*. The EEG electrodes $F5, F6, C3, C4, P5$ and $P6$ were used in the experiments.

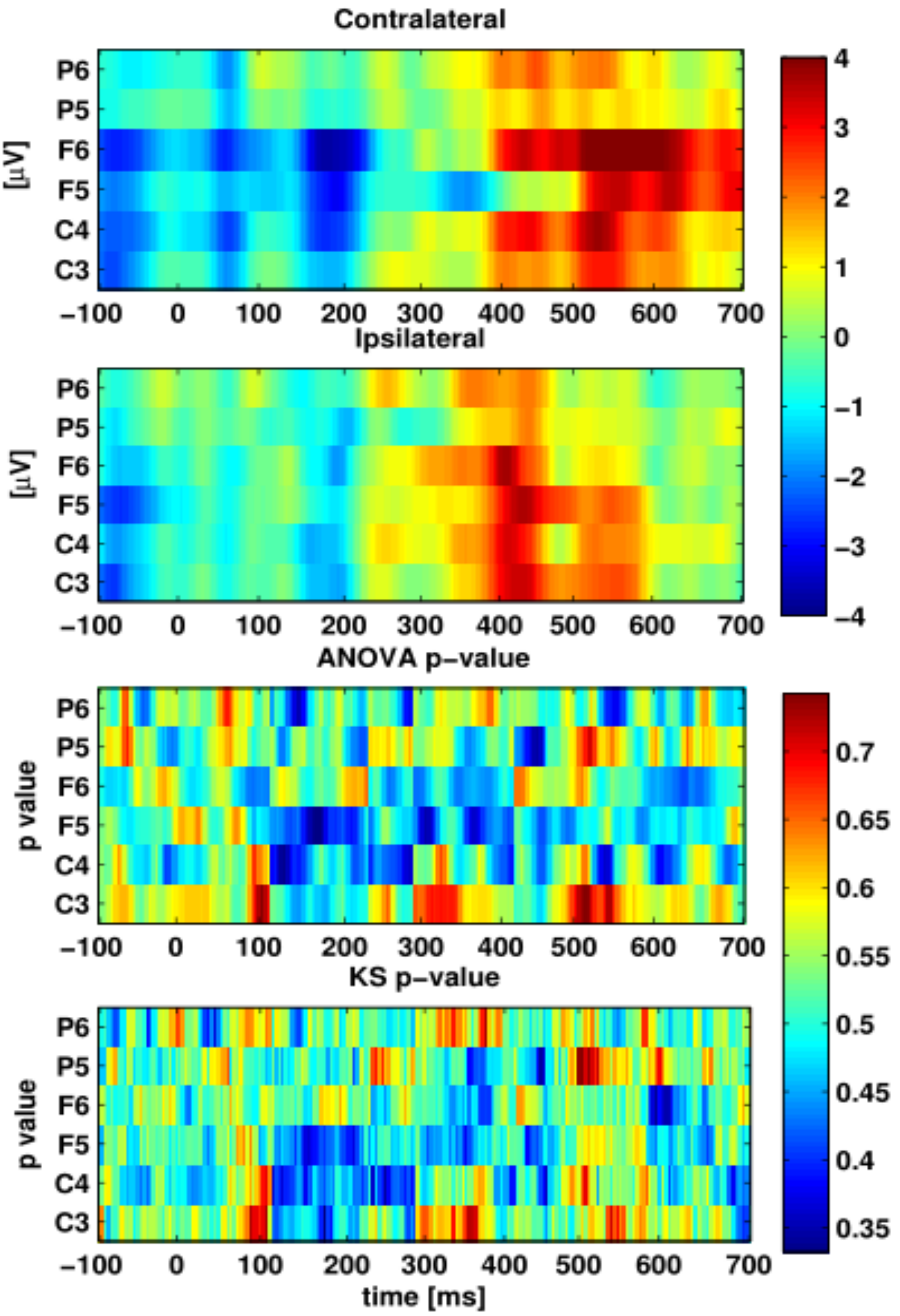

Figure 6.4: Grand mean average ERP responses of all seven subjects and the six electrodes plotted separately in each panel in *contralateral* (top panel) vs. *ipsilateral* (second from the top panel) stimulus direction presentation settings. The significant differences between the both responses can be found, as visualized by the color with p-values of *ANOVA–test* and *KS–test* results (statistical significance for $p < 0.05$) in the third panel and fourth panel, respectively. EEG electrodes $F5, F6, C3, C4, P5$ and $P6$ were used in the experiment.

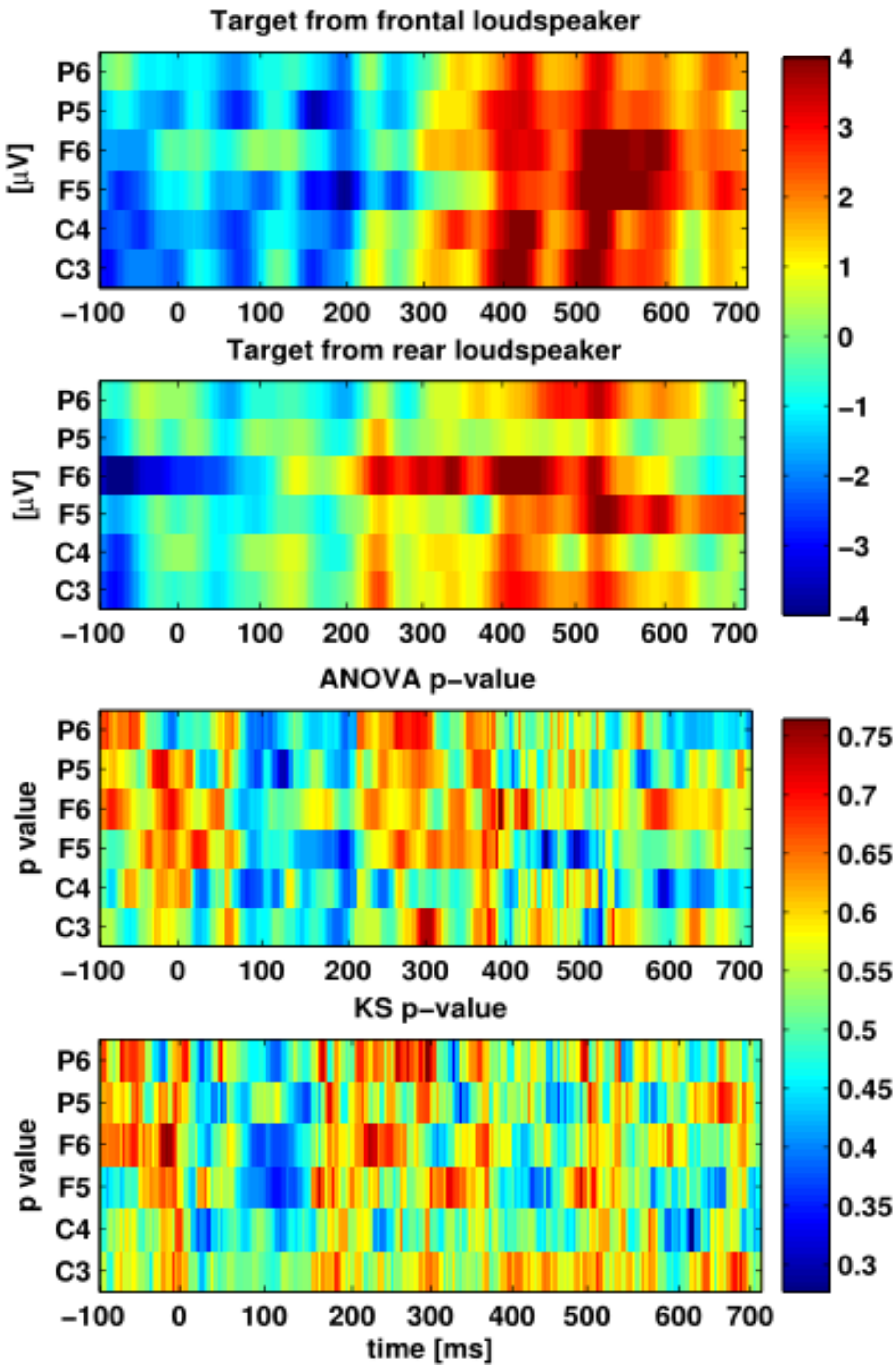

FIGURE 6.5: Grand mean averaged ERP responses of all subjects and the six electrodes plotted separately in each panel for *the frontal–* (top panel) and *the rear–loudspeaker* (second from the top panel) ERP responses. The significant differences between the both responses can be found, as visualized by the color with p-values of *ANOVA–test* and *KS–test* results (statistical significance at $p < 0.05$) in the third and fourth from the top panelz respectively. EEG electrodes $F5, F6, C3, C4, P5$ and $P6$ were used in the experiment.

case) $\mathbf{x} = [x_1, x_2, \ldots, x_l]^T$ based on probability maximization to the class

$$\omega_m = \arg\max_{\omega_i} \prod_{j=1}^{l} p(x_j|\omega_i), \quad i = 1, 2, \ldots, M, \tag{6.1}$$

with an assumption that the individual features x_j, $j = 1, 2, \ldots, l$, shall be statistically independent. It turns out that the NBC can be very robust also to violations of the independence assumption [34].

The results of the successful application of the NBC technique are presented in the next section.

6.4 Results

As the outcome of the presented research I have obtained the results showing that for the both experimental settings of saBCI offline paradigm the classical $P300$ latency could be improved with the $N2ac$ and $N2fr$ features identified with the $p-$values calculated using the ANOVA or KS–tests for significance. I summarize below the obtained results.

6.4.1 The Classification Results from the $P300$ ERP Latencies in the Classical Oddball Paradigm Setting

The first summary of classification results is presented in Table 6.1, where classification accuracies for the features drawn from $P300$ are shown. All of the subjects performed above the chance level of 50% for single feature latencies of $P300$ based on ANOVA– and KS–tests $p-$values. The proposed "hand–picked" feature set identification using the ANOVA– and KS–tests of significant ERP samples allowed us to boost the classification results (in both cases with the maximum classification boost of 29%) using the *leave–one–out* cross validation [34] for the NBC technique.

6.4.2 The Classification Results from the $N2ac$ ERP Feature in the Ipsilateral vs. Contralateral Settings

The results of the approach to compare *ipsilateral* and *contralateral* to target evoked potentials have been summarized in the Table 6.2, based on the ERP features drawn from results of the ANOVA– and KS–tests analysis are summarized in Figure 6.4. The classification accuracy results have been 22% boosted in the best case with ANOVA–test based feature extraction. For KS–test based case the classification accuracy results have been even 28% enhanced in the best. In the both enhancement cases the same method of the NBC *leave–one–out* cross validation classification was applied.

6.4.3 The classification Results of the $N2fr$ ERP Feature Extraction Method

The classification results of the proposed approach to compare *frontal–* and *rear–loudspeaker* originating target stimuli have been summarized in Table 6.3, based on the ERP features selection from results of the ANOVA– and KS–tests analysis as summarized in Figure 6.5. The classification accuracy results have been also in this case 26% boosted in the best cases.

6.4.4 Analysis of Information Transfer Rate Improvement Results

The ITR results are summarized in Tables 6.4, 6.5, and 6.6 respectively. For the all cases of the $P300$, the $N2ac$, and the $N2fr$, there were significant increases of the ITRs for the majority of subjects.

6.5 Conclusions

In this chapter I presented three approaches leading to improvement of the classification accuracy and ITR in the offline saBCI paradigm. The improvement was

TABLE 6.1: The classification results for ERP latencies in $P300$ responses for *target* vs. *non–target* paradigm. The three sets (whole ERP, $P300$ latencies optimized by ANOVA mehtod and KS method) classification results are compared.

subject	noise stimulus type	conventional whole ERP [%]	ANOVA $P300$ [%]	KS $P300$ [%]
#1	pink	61	71	71
	white	65	68	67
#2	pink	61	85	78
	white	63	72	71
#3	pink	58	62	65
	white	55	59	67
#4	pink	49	68	53
	white	55	84	79
#5	pink	48	69	82
	white	56	83	81
#6	pink	55	70	66
	white	57	75	64
#7	pink	54	82	83
	white	65	81	72

obtained from introducing the novel ERP feature extraction methods in the $P300$ latency, the $N2ac$ and the $N2fr$.

The first improvement was presented in form of classification rates comparison for the three ERP feature sets of the $P300$ latencies processed separately with the ANOVA– and KS–tests, versus the classical whole ERP features. The classification accuracy was increased for all the subjects at maximum of 29% boost (ITR improvement at maximum of 34 bit/min). This is a very good result giving a possibility to further improve the auditory based BCI paradigm.

The second improvement step was based on the $N2ac$ concept. The obtained classification and ITR improvement are also very encouraging.

The third improvement step was based on the $N2fr$ responses in the front–back to the subject head stimulus presentation. The classification and ITR improvement confirm that our new finding could improve the final classification results.

Table 6.2: The classification results for ERP latencies in $N200$ responses for *ipsilateral* vs. *contralateral* paradigm. The three feature sets (whole ERP (0 ms – 700 ms), $N200$ latencies optimized by ANOVA method and KS method) classification results are compared.

subject	noise stimulus type	conventional whole ERP [%]	ANOVA $N200$ [%]	KS $N200$ [%]
#1	pink	51	58	57
	white	35	51	56
#2	pink	46	57	58
	white	38	64	59
#3	pink	51	61	57
	white	53	60	65
#4	pink	52	67	59
	white	51	67	71
#5	pink	55	60	58
	white	38	60	52
#6	pink	41	51	52
	white	43	48	57
#7	pink	53	57	65
	white	45	65	73

The three main achievements reported in the chapter allowed us to boost accuracy of the novel saBCI paradigm in offline mode which is a step forward in the non–vision based interfacing strategies.

The obtained results reveal that not only the cortical auditory information processing centers related to the cognitive streams could be utilized for BCI purposes. Also the differences in ERPs at early latencies before 300 ms are useful and they guarantee good classification results and resulting ITR scores. These results reveal that the very early spatial auditory ERPs are potentially interesting for faster BCI applications.

TABLE 6.3: The classification results for ERP latencies in $N200$ responses for *targets from frontal loudspeaker* vs. *targets from rear loudspeaker* paradigm. The three feature sets (whole ERP (0 ms – 700 ms), $N200$ latencies optimized by ANOVA method and KS method) classification results are compared.

subject	noise stimulus type	conventional whole ERP [%]	ANOVA $N200$ [%]	KS $N200$ [%]
#1	pink	46	52	50
	white	51	63	77
#2	pink	40	66	53
	white	36	51	56
#3	pink	52	65	58
	white	51	55	67
#4	pink	63	73	70
	white	53	60	57
#5	pink	43	57	56
	white	46	57	51
#6	pink	37	56	58
	white	52	62	56
#7	pink	57	68	65
	white	44	51	55

TABLE 6.4: The ITR of *target* vs. *non-target*, for the three ERP sets related classification approaches using *wholeERP* , $P300$ extracted with ANOVA and KS.

subject	noise stimulus type	Whole ERP [bit/min]	ANOVA $P300$ [bit/min]	KS $P300$ [bit/min]
#1	pink	3.25	13.13	13.13
	white	6.59	9.56	8.51
#2	pink	3.52	39.02	23.98
	white	4.93	14.45	13.13
#3	pink	1.85	4.19	6.59
	white	0.72	2.35	8.51
#4	pink	0.00	9.56	0.26
	white	0.72	36.57	25.85
#5	pink	0.00	10.68	31.99
	white	1.04	34.23	29.85
#6	pink	0.72	11.87	7.52
	white	1.42	18.87	5.73
#7	pink	0.46	31.99	34.23
	white	6.59	29.85	14.45

Table 6.5: The ITR of N2ac, for the three ERP sets related classification approaches using *wholeERP* , $N200$ extracted with ANOVA and KS.

subject	noise stimulus type	Whole ERP [bit/min]	ANOVA $N200$ [bit/min]	KS $N200$ [bit/min]
#1	pink	0.03	1.85	1.42
	white	0.00	0.03	1.04
#2	pink	0.00	1.42	1.85
	white	0.00	5.73	2.35
#3	pink	0.03	3.52	1.42
	white	0.26	2.90	6.59
#4	pink	0.12	8.51	2.35
	white	0.03	8.51	13.13
#5	pink	0.72	2.90	1.85
	white	0.00	2.90	0.12
#6	pink	0.00	0.03	0.12
	white	0.00	0.00	1.42
#7	pink	0.26	1.42	6.59
	white	0.00	6.59	15.85

Table 6.6: The ITR of *targets* from front loudspeaker and *targets* from rear loudspeaker, for the three ERP sets related classification approaches using *wholeERP* , $N200$ extracted with ANOVA and KS.

subject	noise stimulus type	Whole ERP [bit/min]	ANOVA $N200$ [bit/min]	KS $N200$ [bit/min]
#1	pink	0.00	0.12	0.00
	white	0.03	4.93	22.19
#2	pink	0.00	7.52	0.26
	white	0.00	0.03	0.72
#3	pink	0.12	6.59	1.85
	white	0.03	0.72	8.51
#4	pink	4.93	15.85	11.87
	white	0.26	2.90	1.41
#5	pink	0.00	1.42	1.04
	white	0.00	1.42	0.03
#6	pink	0.00	1.04	1.85
	white	0.12	4.19	1.04
#7	pink	1.42	9.56	6.59
	white	0.00	0.03	0.72

Chapter 7

Conclusion of thesis

7.1 Summary of thesis

In this thesis, I designed a novel auditory BCI based on combined sound timbre and horizontal plane spatial locations as informative cues. I also studied EEG signal processing and classification techniques in order to improve the BCI performance and use them in a novel spatial auditory BCI system, with a major objectives of enhancement of ERP classification results.

In order to achieve the major objective, I have first proposed "EEG channel selection" and ERP classification, with fewer channels more quickly and more effectively improve the BCI performance. Concerning "EEG channel selection", I have proposed to use a method name ROC, which is based on quantify the separability of two single-trial response distributions. Such method is related in a direct and natural way to cost/benefit analysis of diagnostic decision making. Concerning ERP classification, I proposed to use LDA method which used in statistics, pattern recognition and machine learning to find a linear combination of features which characterizes or separates two or more classes of objects or events. The resulting combination may be used as a linear classifier or more commonly for dimensionality reduction before later classification (see chapter 3). These algorithm can learn and extract the discriminant features, such features corresponding to the brain activity in Regions of Interest (ROI). Our evaluations suggested that the ROI and "channels selection" can improve the ERP classification accuracy.

Concerning feature extraction, I propose to use a method (ANOVA: analysis of various) for the extraction of discriminative features in electroencephalography (EEG) evoked potential latency ($P300$ response). This method is a collection of statistical models used to analyze the differences between group means (in our case are *target* and *non-target*) and their associated procedures (such as "variation" among and between groups). In ANOVA setting, the observed variance in a particular variable is partitioned into components attributable to different ERP responses of variation. In its simplest form, ANOVA provides a statistical test of whether or not the means of *target* and *non-target* are different or not. More particularly, a naive Bayesian classifier for classifier the subject's chosen targets and ignored non-targets. A naive Bayes classifier is a simple probabilistic classifier based on applying Bayes' theorem with strong (naive) independence assumptions. However, GNBC classifier yielded similar or even better results on our experimental data than linear discrimination analysis methods. Our results reveal that the feature extraction based on ANOVA and ERP classification with GNBC can obtain a good results in spatial auditory BCI system (see chapter 4).

Chapter 4 BCI paradigms benefit mostly from the $P300$ ERP latencies. I show the further enhancement of the classification results in spatial auditory paradigms by incorporating the $N200$ latencies, which differentiate the brain responses to lateral, in relation to the subject head, sound locations in the auditory space. In order to extract the $P300$ and $N200$ responses, I proposed t-test, which can be used to determine if two sets of data (in our case *target* vs. *non-target*; "ipsilateral" vs "contralateral") are significantly different from each other. The results reveal that those early spatial auditory ERPs boost online classification results of the BCI application. The online BCI experiments with the multi–command BCI prototype support our research hypothesis could improve classification accuracies and information–transfer–rates.

In chapter 6, I presented three approaches leading to improvement of the classification accuracy and ITR in the offline saBCI paradigm. The improvement was obtained from introducing the novel ERP feature extraction methods in the $P300$ latency, the $N2ac$ and the $N2fr$. The first improvement was presented in form of classification rates comparison for the three ERP feature sets of the $P300$ latencies processed separately with the ANOVA– and KS–tests, versus the classical whole ERP features. The classification accuracy was increased for all the subjects at maximum of 29% boost (ITR improvement at maximum of 34 bit/min). This is a very good result giving a possibility to further improve the auditory based BCI paradigm. The second improvement step was based on the $N2ac$ concept. The obtained classification and ITR improvement are also very encouraging. The third improvement step was based on the $N2fr$ responses in the front–back to the subject head stimulus presentation. The classification and ITR improvement confirm that our new finding could improve the final classification results. The three main achievements reported in the chapter allowed us to boost accuracy of the novel saBCI paradigm in offline mode which is a step forward in the non–vision based interfacing strategies. The obtained results reveal that not only the cortical auditory information processing centers related to the cognitive streams could be utilized for BCI purposes. Also the differences in ERPs at early latencies before 300 ms are useful and they guarantee good classification results and resulting ITR scores. These results reveal that the very early spatial auditory ERPs are potentially interesting for faster BCI applications.

7.2 Discussion

In the chapter it has been shown that in contrary to the contemporary results with the spatial auditory BCI paradigms, which fail to utilize rear-head loudspeakers, it is possible to achieve good results for a fully surround sound octagonal loudspeakers setup. I have also shown that various spatial and timber sound sources generate different ERP latencies ($P300$, $N200$, $N2fr$, $N2ac$) and amplitude responses together with naive Bays classifier possible to utilize in novel spatial auditory BCI paradigms. These are the very encouraging results, providing the possibility further to improve the auditory paradigm based BCI. The main achievement reported in the thesis allows us to improve the spatial aBCI paradigm in the offline mode, which is a step forward in non-vision based interfacing strategies. I have also shown that, in comparison with contemporary applications of spatial auditory BCI paradigms which fail to utilize rear–to–the–head loudspeakers, it is possible to utilize all spatial horizontal sound directions thanks to the proposed classification improvement approach based on the "hand–picked" ERP latencies.

Bibliography

[1] Jonathan Wolpaw and Elizabeth Winter Wolpaw, editors. *"Brain-Computer Interfaces: Principles and Practice"*. Oxford University Press, 2012.

[2] P.Sajda, K.R. Mueller, and K.V Shenoy. "brain computer interfaces". In *IEEE signal Processing Magazine*, January 2008.

[3] L. Farwell and E. Donchin. "talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials". In *Electroencephalogr. Clin. Neurophysiol*, pages 510–523, 1988.

[4] J. E. Huggins, P. A. Wren, and K. L. Gruis. "what would brain computer interface users want? opinions and priorities of potential users with amyotrophic lateral sclerosis". In *Amyotroph.Lateral Scler.*, 2009. doi: 10.3109/17482968.

[5] A.R. Murguialday, J. Hill, M. Bensch, S. Martens, S. Halder, F. Nijboer, B. Scholkopf, N. Birbaumer, and A. Gharabaghi. In *Clin,Neurophysiol.*, pages 925–933.

[6] P. Brunner, S. Joshi, S. Briskin, J. R. Wolpaw, H. Bischof, and G. Schalk. "does the "p300" speller depend on eye gaze ?". In *J.Neural Eng.7*, 2010.

[7] M. S. Treder and B. Blankertz. "covert attention and visual speller design in an erp-based brain computer interface". In *Behav. Brain Funct*, 2010.

[8] N. Hill, T. Lal, K. Bierig, N. Birbaumer, and B. Scholkopf. "an auditory paradigm for brain computer interfaces". In *Advances in Neural Infromation Processing Systems*, pages 569–576, 2005.

[9] S. Halder, M. Rea, R. Andreoni, F. Nijboer, E.M. Hammer, S.C. Kleih, N. Birbaumer, and A. Kubler. "an auditory oddball brain computer interface for binary choices". In *Clin. Neurophysiol.*, pages 516–523, 2010.

[10] M. Schreuder, M. Tangeermann, and B. Blankertz. "initial results of a high speed spatial auditory bci". In *Int.J.Bioelectromagn*, pages 105–109, 2009.

[11] J. Hohne, M. Schreuder, B. Blankertz, and M. Tangermann. "two dimensional auditory p300 speller with predictive text system". In *Conf. Proc. IEEE Eng. Med. Biol. Soc. 1*, pages 4185–4188, 2010.

[12] M. Schreduer, B. Blankertz, and M. Tangermann. "a new auditory multi-class brain-computer interface paradigm: Spatial hearing as an informative cue",. In *PLoS ONE 5 e9813*, 2010. doi: 10.1371/journal.pone.0009813.

[13] A. Belitski, J. Farquhar, and P. esain. "p300 audio-visual speller". In *J. Neural Eng. 8*, 2011.

[14] E. Niedermeyer and F. L. Da Silva, editors. *"Electroencephalography: Basic Principles, Clinical Applications, and Related Fields"*. Lippincott Williams & Wilkins, 5 edition, 2004.

[15] J. Chapin, K. Moxon, R. Markowitz, and M. Nicolelis. "real-time control of a robot arm using simultaneously recordedc neurons in the motor cortex". In *Nature Neuroscience*, pages 664–670, 1999.

[16] N. K. Logothetis, J. Pauls, M. Augath, T. Trinath, and A. Oeltermann. "neurophysiological investigation of the basis of the fmri signal,". In *Nature Neuroscience*, pages 150–157, 2001.

[17] V. A. and B. Chance. "non-invasive optical spectroscopy and imaging of human brain function,". In *Trends in Neuroscience*, pages 435–442, 1997.

[18] D. Cohen. "magnetoencephalography: evidence of magnetic fields produced by alpha rhythm currents,". In *Science*, pages 784–786, 1968.

[19] D. Cohen. "magnetoencephalography: detection of the brain's electrical activity with a superconducting magnetometer,". In *Science*, pages 664–666, 1972.

[20] H. Berger. "uber das electrenkephalogramm des menchen,". In *Arch Psychiat*, 1929.

[21] J. L. Steven. "an introduction to the event-related potential technique,". In *The MIT Press*, 2005.

[22] L. A. Farwell and E. Donchin. "talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials,". In *Electroencephalography and Clinical Neurophysiology*, pages 510–523, 1988.

[23] E. W. Sellers, A. Kubler, and E. Donchin. "brain-computer interface research at the university of south florida cognitive psychophysiology laboratory: the p300 speller". In *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, pages 221–214.

[24] T. M. Rutkowski, A. Cichocki, and D. P. Mandic. "spatial auditory paradigms for brain computer/machine interfacing". In *International Workshop On The Principles and Applications of Spatial Hearing 2009 (IWPASH 2009) - Proceedings of the International Workshop*, page P5, Miyagi-Zao Royal Hotel, Sendai, Japan, November 11-13, 2009. doi: 10.1142/9789814299312_0025. URL http://dx.doi.org/10.1142/9789814299312_0025.

[25] T. M. Rutkowski, T. Tanaka, Q. Zhao, and A. Cichocki. "spatial auditory bci/bmi paradigm - multichannel emd approach to brain responses estimation". In *In proceedings of the Second APSIPA Annual Summit and Conference (APSIPA ASC 2010)*, pages 197–202, 2010.

[26] J. Lewald, K. A. J. Riederer, T. Lentz, and I. G. Meister. "processing of sound localization in human cotex". In *European Journal of Neuroscience*, pages 1261–1270, 2008.

[27] J. Schnupp, I. Nelken, and A. King. *"Auditory Neuroscience - Making Sense of Sound"*. MIT Press, 2010.

[28] J. C. Middlebrooks and D. M.Green. "sound localization by human listeners". In *Annu Rev Psychol 42:*, pages 135–159.

[29] D. S. Brungart, N. I. Durlach, and W. M. Rabinowitz. "auditory localization of nearby sources",. In *J Acoust Soc Am 106:*, pages 1956–1968.

[30] T. A. Mondor and R. J. Zatorre. "shifting and focusing auditory spatial attention". In *J Exp Psychol Hum Percept Perform 21:*, pages 387–409.

[31] W. A. Teder-Salejarvi and S. A. Hillyard. "the gradient of spatial auditory attention in free field: an event-related potential study". In *Percept Psychophys 60:*, pages 1228–1242.

[32] C. Gugerl, S. Daban, E. Sellers, C. Holzner, G. Krausz, R. Carabalona, F.Gramatica, and G. Edlinger. "how many people are able to control a p300-based brain computer interface (bci)?". In *Neuroscience Letters*, pages 94–98.

[33] Wellcome Trust Centre for Neuroimaging. "statistical parametric mapping - spm8 package".

[34] T. Sergios and K. Konstantinos. *"Pattern Recognition"*. Acedemic Press, fourh edition, 2009.

[35] T. Fawcett. "an introduction to roc analysis". In *Pattern Recognition Letters*, pages 861–874, 2006.

[36] R Development Core Team. "r: A language and environment for statistical computing". In *R Foundation for Statistical Computing, Vienna, Austria*, 2009.

[37] J. Tuszynski. "catools: Tools: moving window statistics". In *GIF, Base64, ROC AUC, etc*, 2011.

[38] Z. Cai, H. Terasawa, S. Makino, T. Yamada, and T. M. Rutkowski. "sound timbre and spatial location as informative cues in auditory BCI - brain evoked potential enhancement and involuntary eye movements artifacts suppression approach". In *Proceedings of the Third APSIPA Annual Summit and Conference (APSIPA ASC 2011)*, page paper 241 (6 pages), Xi'an, China, October 18 - 21, 2011. APSIPA.

[39] J. Valer, T. Daisuke, and D. Ippeita. "10/20, 10/10, and 10/5 systems revisited: Their validity as relative head-surface-based positioning systems". *NeuroImage*, 34(4):1600 – 1611, 2007. ISSN 1053-8119. doi: 10.1016/j.neuroimage.2006.09.024. URL http://www.sciencedirect.com/science/article/pii/S1053811906009724.

[40] T. M. Rutkowski. "auditory brain-computer/machine interface paradigms design". In Eric Cooper, Victor Kryssanov, Hitoshi Ogawa, and Stephen Brewster, editors, *Haptic and Audio Interaction Design*, volume 6851 of *Lecture Notes in Computer Science*, pages 110–119. Springer Berlin / Heidelberg, 2011. ISBN 978-3-642-22949-7. doi: 10.1007/978-3-642-22950-3_12. URL http://dx.doi.org/10.1007/978-3-642-22950-3_12.

[41] H. Johannes, S. Martijn, B. Benjamin, and T. Michael. "a novel 9-class auditory ERP paradigm driving a predictive text entry system". *Frontiers in Neuroscience*, 5:Article 99, 2011. doi: 10.3389/fnins.2011.00099. URL http://www.frontiersin.org/Journal/Abstract.aspx?s=763&name=neuroprosthetics&ART_DOI=10.3389/fnins.2011.00099.

[42] M. Schreuder, T. Rost, and M. Tangermann. "listen, you are writing! Speeding up online spelling with a dynamic auditory BCI". *Frontiers in Neuroscience*, 5:Article 112, 2011. doi: 10.3389/fnins.2011.00112. URL http://www.frontiersin.org/Journal/Abstract.aspx?s=763&name=neuroprosthetics&ART_DOI=10.3389/fnins.2011.00112.

[43] M. Schreuder, B. Blankertz, and M. Tangermann. "a new auditory multi-class brain-computer interface paradigm: Spatial hearing as an informative cue". *PLoS ONE*, 5(4):e9813, 04 2010. doi: 10.1371/journal.pone.0009813. URL http://dx.doi.org/10.1371%2Fjournal.pone.0009813.

[44] Z. Cai, S. Makino, T. Yamada, and T. M. Rutkowski. "multichannel analysis of ERP at early and p300 latencies in application for spatial auditory bci improvement". In *Proceedings of the Fouth APSIPA Annual Summit and Conference (APSIPA ASC 2012)*, page paper 355 (7 pages), Hollywood, America, December 3 - 6, 2012. APSIPA.

[45] T. M. Mitchell. "machine learning". In *New York:McGraw-Hill*, 1997.

[46] A. Schloegl, J. Kronegg, J.E. Huggins, and S. G. Mason. Evaluation criteria in BCI research. In G. Dornhege, J. R. Millan, T. Hinterberger, and K. R.Mueller D. J. McFarland, editors, *Towards Brain-Computer Interfacing*, pages 327–342. MIT Press, 2007.

[47] Z. Cai, S. Makino, and T. M. Rutkowski. "brain evoked potential latencies optimization for spatial auditory brain computer interface". *Cognitive Computation*, 5(2), 2013. ISSN 1866-9956. doi: 10.1007/s12559-013-9228-x.

[48] M. L. Gamble and S. J. Luck. "N2ac: An ERP component associated with the focusing of attention within an auditory scene". *Psychophysiology*, 48:1057–1068, 2011. doi: 10.3389/fnins.2011.00099. URL http://www.frontiersin.org/Journal/Abstract.aspx?s=763&name=neuroprosthetics&ART_DOI=10.3389/fnins.2011.00099.

[49] Helmut F. van Emden. *"Statistics for Terrified Biologists"*. Blackwell Publishing, 2008.

Author's Paper List

1. Z.Cai, N.Kitawaki, T.Yamada, "Comparison of the MOS Evaluation of Chinese and Japanese," IEICE, B-11-12,p.443 Mar. 2009.

2. Z.Cai, N.Kitawaki, T.Yamada,"Comparison of MOS Evaluation Characteristics by Chinese, Japanese, and English in IP Telephony," IEICE Technical Report CQ2009-28, P.17-20 Sep. 2009.

3. Z.Cai, N. Kitawaki, T. Yamada, and S. Makino, "Comparison of MOS evaluation characteristics for Chinese, Japanese, and English in IP telephony," Proc. International Universal Communication Symposium, IUCS 2010, pp. 112 - 115, Oct. 2010.

4. T. M. Rutkowski, Q. Zhao, D. P. Mandic, Z. Cai, A. Cichocki, S. Makino, and A. W. Przybyszewski, "New EEG components separation method: Data driven Huang-Hilbert transform application to auditory BMI paradigm," in Neuroscience 2011 Abstracts, p.627.15/AAA32, Society for Neuroscience, Nov. 2011.

5. Z. Cai, H. Terasawa, S. Makino, T. Yamada, and T. M. Rutkowski,@"Spatial location and sound timbre as informative cues in auditory BCI/BMI - electrodes position optimization for brain evoked potential enhancement," Proc. APSIPA ASC 2011, p.paper.241, Oct. 2011

6. Z. Cai, S. Makino, and T. M. Rutkowski, gSpatial auditory BCI/BMI paradigm,h in Proceedings of The Second International Workshop on Brain Inspired Computing (BIC 2012), p. poster.2.1, NIMS, June 2012.

7. T. M. Rutkowski, H. Mori, Y. Matsumoto, Z. Cai, M. Chang, N. Nishikawa, S. Makino, and K. Mori, "Haptic BCI paradigm based on somatosensory evoked potential," *Arxiv preprint arXiv:1207.5720*, Oct. 2012.

8. T. M. Rutkowski, Z. Cai, N. Nishikawa, Y. Matsumoto, S. Makino, D. Looney, D. Mandic, Z. Struzik, and A. W. Przybyszewski, gA multi-command spatial auditory BMI based on evoked EEG responses from real and virtual sound stimuli,h in Neuroscience 2012 Abstracts, no. 891.16/NN4 in Brain Machine Interfaces: Non-Invasive Methods, p.abstract.12387, SfN, Oct. 2012.

9. M. Chang, N. Nishikawa, Z. Cai, S. Makino, and T. M. Rutkowski, gPsychophysical responses comparison in spatial auditory, visual, and audiovisual BCI-spelling paradigms,h in Proceedings of The 6th International Conference on Soft Computing and Intelligent Systems, and The 13th International Symposium on Advanced Intelligent Systems, p.2154 - 2157, Nov. 2012.

10. Z. Cai, S. Makino, T. Yamada and T. M. Rutkowski, gspatial auditory BCI paradigm utilizing N200 and P300,h in Proceedings of the Fourth APSIPA Annual Summit and Conference (APSIPA ASC 2012), p.paper.355, Dec. 2012

11. Z. Cai, S. Makino, and T. M. Rutkowski. hbrain evoked potential latencies optimization for spatial auditory brain computer interfaceh. Cognitive Computation, 5(2), ISSN 1866-9956. doi: 10.1007/s12559-013-9228-x, 2013.

12. Z. Cai, S. Makino and T. M. Rutkowski. gSpatial Auditory BCI with ERP Responses to Front?Back to the Head Stimuli Distinction Support,h in Proceedings of the Fifth APSIPA Annual Summit and Conference (APSIPA ASC 2013), p.paper.85, Oct, 2013.